

SEQUENCE LISTING



(1) GENERAL INFORMATION:

- (i) APPLICANT: ARTHUR, MICHEL
DUKTA-MALEN, SYLVIE
MOLINAS, CATHERINE
COURVALIN, PATRICE
- (ii) TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
- (iii) NUMBER OF SEQUENCES: 54
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
P.C.
 - (B) STREET: 1755 S. Jefferson Davis Highway, Suite 400
 - (C) CITY: Arlington
 - (D) STATE: Virginia
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 22202
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/286,819
 - (B) FILING DATE: 05-AUG-1994
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/174,682
 - (B) FILING DATE: 28-DEC-1993
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/917,146
 - (B) FILING DATE: 10-AUG-1992
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/FR/91/00855
 - (B) FILING DATE: 29-OCT-1991
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: FR 9013579
 - (B) FILING DATE: 31-OCT-1990
 - (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
 (A) NAME: Oblon, Norman F.
 (B) REGISTRATION NUMBER: 24,618
 (C) REFERENCE/DOCKET NUMBER: 660-060-0 PCT

(ix) TELECOMMUNICATION INFORMATION:
 (A) TELEPHONE: (703) 413-3000
 (B) TELEFAX: (703) 413-2220
 (C) TELEX: 248855 OPAT UR

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 966 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..966

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG AAT AAC ATC GGC ATT ACT GTT TAT GGA TGT GAG CAG GAT GAG GCA	48
Met Asn Asn Ile Gly Ile Thr Val Tyr Gly Cys Glu Gln Asp Glu Ala	
1 5 10 15	
GAT GCA TTC CAT GCT CTT TCG CCT CGC TTT GGC GTT ATG GCA ACG ATA	96
Asp Ala Phe His Ala Leu Ser Pro Arg Phe Gly Val Met Ala Thr Ile	
20 25 30	
ATT AAC GCC AAC GTG TCG GAA TCC AAC GCC AAA TCC GCG CCT TTC AAT	144
Ile Asn Ala Asn Val Ser Glu Ser Asn Ala Lys Ser Ala Pro Phe Asn	
35 40 45	
CAA TGT ATC AGT GTG GGA CAT AAA TCA GAG ATT TCC GCC TCT ATT CTT	192
Gln Cys Ile Ser Val Gly His Lys Ser Glu Ile Ser Ala Ser Ile Leu	
50 55 60	
CTT GCG CTG AAG AGA GCC GGT GTG AAA TAT ATT TCT ACC CGA AGC ATC	240
Leu Ala Leu Lys Arg Ala Gly Val Lys Tyr Ile Ser Thr Arg Ser Ile	
65 70 75 80	
GGC TGC AAT CAT ATA GAT ACA ACT GCT GCT AAG AGA ATG GGC ATC ACT	288
Gly Cys Asn His Ile Asp Thr Thr Ala Ala Lys Arg Met Gly Ile Thr	
85 90 95	
GTC GAC AAT GTG GCG TAC TCG CCG GAT AGC GTT GCC GAT TAT ACT ATG	336
Val Asp Asn Val Ala Tyr Ser Pro Asp Ser Val Ala Asp Tyr Thr Met	
100 105 110	

ATG	CTA	ATT	CTT	ATG	GCA	GTA	CGC	AAC	GTA	AAA	TCG	ATT	GTG	CGC	TCT	384
Met	Leu	Ile	Leu	Met	Ala	Val	Arg	Asn	Val	Lys	Ser	Ile	Val	Arg	Ser	
		115					120									
GTG	GAA	AAA	CAT	GAT	TTC	AGG	TTG	GAC	AGC	GAC	CGT	GGC	AAG	GTA	CTC	432
Val	Glu	Lys	His	Asp	Phe	Arg	Leu	Asp	Ser	Asp	Arg	Gly	Lys	Val	Leu	
		130				135					140					
AGC	GAC	ATG	ACA	GTT	GGT	GTG	GTG	GGA	ACG	GGC	CAG	ATA	GGC	AAA	GCG	480
Ser	Asp	Met	Thr	Val	Gly	Val	Val	Gly	Thr	Gly	Gln	Ile	Gly	Lys	Ala	
					150					155					160	
GTT	ATT	GAG	CGG	CTG	CGA	GGA	TTT	GGA	TGT	AAA	GTG	TTG	GCT	TAT	AGT	528
Val	Ile	Glu	Arg	Leu	Arg	Gly	Phe	Gly	Cys	Lys	Val	Leu	Ala	Tyr	Ser	
				165					170					175		
CGC	AGC	CGA	AGT	ATA	GAG	GTA	AAC	TAT	GTA	CCG	TTT	GAT	GAG	TTG	CTG	576
Arg	Ser	Arg	Ser	Ile	Glu	Val	Asn	Tyr	Val	Pro	Phe	Asp	Glu	Leu	Leu	
			180					185					190			
CAA	AAT	AGC	GAT	ATC	GTT	ACG	CTT	CAT	GTG	CCG	CTC	AAT	ACG	GAT	ACG	624
Gln	Asn	Ser	Asp	Ile	Val	Thr	Leu	His	Val	Pro	Leu	Asn	Thr	Asp	Thr	
		195					200					205				
CAC	TAT	ATT	ATC	AGC	CAC	GAA	CAA	ATA	CAG	AGA	ATG	AAG	CAA	GGA	GCA	672
His	Tyr	Ile	Ile	Ser	His	Glu	Gln	Ile	Gln	Arg	Met	Lys	Gln	Gly	Ala	
		210				215					220					
TTT	CTT	ATC	AAT	ACT	GGG	CGC	GGT	CCA	CTT	GTA	GAT	ACC	TAT	GAG	TTG	720
Phe	Leu	Ile	Asn	Thr	Gly	Arg	Gly	Pro	Leu	Val	Asp	Thr	Tyr	Glu	Leu	
					230					235					240	
GTT	AAA	GCA	TTA	GAA	AAC	GGG	AAA	CTG	GGC	GGT	GCC	GCA	TTG	GAT	GTA	768
Val	Lys	Ala	Leu	Glu	Asn	Gly	Lys	Leu	Gly	Gly	Ala	Ala	Leu	Asp	Val	
				245					250					255		
TTG	GAA	GGA	GAG	GAA	GAG	TTT	TTC	TAC	TCT	GAT	TGC	ACC	CAA	AAA	CCA	816
Leu	Glu	Gly	Glu	Glu	Glu	Phe	Phe	Tyr	Ser	Asp	Cys	Thr	Gln	Lys	Pro	
			260					265					270			
ATT	GAT	AAT	CAA	TTT	TTA	CTT	AAA	CTT	CAA	AGA	ATG	CCT	AAC	GTG	ATA	864
Ile	Asp	Asn	Gln	Phe	Leu	Leu	Lys	Leu	Gln	Arg	Met	Pro	Asn	Val	Ile	
		275					280					285				
ATC	ACA	CCG	CAT	ACG	GCC	TAT	TAT	ACC	GAG	CAA	GCG	TTG	CGT	GAT	ACC	912
Ile	Thr	Pro	His	Thr	Ala	Tyr	Tyr	Thr	Glu	Gln	Ala	Leu	Arg	Asp	Thr	
		290				295					300					
GTT	GAA	AAA	ACC	ATT	AAA	AAC	TGT	TTG	GAT	TTT	GAA	AGG	AGA	CAG	GAG	960
Val	Glu	Lys	Thr	Ile	Lys	Asn	Cys	Leu	Asp	Phe	Glu	Arg	Arg	Gln	Glu	
					310					315					320	
CAT	GAA															

✓ (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 322 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asn	Asn	Ile	Gly	Ile	Thr	Val	Tyr	Gly	Cys	Glu	Gln	Asp	Glu	Ala	
1				5					10					15		
Asp	Ala	Phe	His	Ala	Leu	Ser	Pro	Arg	Phe	Gly	Val	Met	Ala	Thr	Ile	
			20					25					30			
Ile	Asn	Ala	Asn	Val	Ser	Glu	Ser	Asn	Ala	Lys	Ser	Ala	Pro	Phe	Asn	
		35					40					45				
Gln	Cys	Ile	Ser	Val	Gly	His	Lys	Ser	Glu	Ile	Ser	Ala	Ser	Ile	Leu	
	50					55					60					
Leu	Ala	Leu	Lys	Arg	Ala	Gly	Val	Lys	Tyr	Ile	Ser	Thr	Arg	Ser	Ile	
65					70					75					80	
Gly	Cys	Asn	His	Ile	Asp	Thr	Thr	Ala	Ala	Lys	Arg	Met	Gly	Ile	Thr	
				85					90					95		
Val	Asp	Asn	Val	Ala	Tyr	Ser	Pro	Asp	Ser	Val	Ala	Asp	Tyr	Thr	Met	
		100						105					110			
Met	Leu	Ile	Leu	Met	Ala	Val	Arg	Asn	Val	Lys	Ser	Ile	Val	Arg	Ser	
		115					120					125				
Val	Glu	Lys	His	Asp	Phe	Arg	Leu	Asp	Ser	Asp	Arg	Gly	Lys	Val	Leu	
	130					135					140					
Ser	Asp	Met	Thr	Val	Gly	Val	Val	Gly	Thr	Gly	Gln	Ile	Gly	Lys	Ala	
145					150					155					160	
Val	Ile	Glu	Arg	Leu	Arg	Gly	Phe	Gly	Cys	Lys	Val	Leu	Ala	Tyr	Ser	
				165					170					175		
Arg	Ser	Arg	Ser	Ile	Glu	Val	Asn	Tyr	Val	Pro	Phe	Asp	Glu	Leu	Leu	
			180					185					190			
Gln	Asn	Ser	Asp	Ile	Val	Thr	Leu	His	Val	Pro	Leu	Asn	Thr	Asp	Thr	
		195					200					205				
His	Tyr	Ile	Ile	Ser	His	Glu	Gln	Ile	Gln	Arg	Met	Lys	Gln	Gly	Ala	
	210					215					220					
Phe	Leu	Ile	Asn	Thr	Gly	Arg	Gly	Pro	Leu	Val	Asp	Thr	Tyr	Glu	Leu	
225					230					235					240	
Val	Lys	Ala	Leu	Glu	Asn	Gly	Lys	Leu	Gly	Gly	Ala	Ala	Leu	Asp	Val	
				245					250					255		

GCT	TTG	CAT	GGC	AAG	TCA	GGT	GAA	GAT	GGA	TCC	ATA	CAA	GGT	CTG	TTT	336
Ala	Leu	His	Gly	Lys	Ser	Gly	Glu	Asp	Gly	Ser	Ile	Gln	Gly	Leu	Phe	
			100					105					110			
GAA	TTG	TCC	GGT	ATC	CCT	TTT	GTA	GGC	TGC	GAT	ATT	CAA	AGC	TCA	GCA	384
Glu	Leu	Ser	Gly	Ile	Pro	Phe	Val	Gly	Cys	Asp	Ile	Gln	Ser	Ser	Ala	
		115					120					125				
ATT	TGT	ATG	GAC	AAA	TCG	TTG	ACA	TAC	ATC	GTT	GCG	AAA	AAT	GCT	GGG	432
Ile	Cys	Met	Asp	Lys	Ser	Leu	Thr	Tyr	Ile	Val	Ala	Lys	Asn	Ala	Gly	
	130					135					140					
ATA	GCT	ACT	CCC	GCC	TTT	TGG	GTT	ATT	AAT	AAA	GAT	GAT	AGG	CCG	GTG	480
Ile	Ala	Thr	Pro	Ala	Phe	Trp	Val	Ile	Asn	Lys	Asp	Asp	Arg	Pro	Val	
145					150					155					160	
GCA	GCT	ACG	TTT	ACC	TAT	CCT	GTT	TTT	GTT	AAG	CCG	GCG	CGT	TCA	GGC	528
Ala	Ala	Thr	Phe	Thr	Tyr	Pro	Val	Phe	Val	Lys	Pro	Ala	Arg	Ser	Gly	
				165					170					175		
TCA	TCC	TTC	GGT	GTG	AAA	AAA	GTC	AAT	AGC	GCG	GAC	GAA	TTG	GAC	TAC	576
Ser	Ser	Phe	Gly	Val	Lys	Lys	Val	Asn	Ser	Ala	Asp	Glu	Leu	Asp	Tyr	
			180					185					190			
GCA	ATT	GAA	TCG	GCA	AGA	CAA	TAT	GAC	AGC	AAA	ATC	TTA	ATT	GAG	CAG	624
Ala	Ile	Glu	Ser	Ala	Arg	Gln	Tyr	Asp	Ser	Lys	Ile	Leu	Ile	Glu	Gln	
		195					200					205				
GCT	GTT	TCG	GGC	TGT	GAG	GTC	GGT	TGT	GCG	GTA	TTG	GGA	AAC	AGT	GCC	672
Ala	Val	Ser	Gly	Cys	Glu	Val	Gly	Cys	Ala	Val	Leu	Gly	Asn	Ser	Ala	
	210					215					220					
GCG	TTA	GTT	GTT	GGC	GAG	GTG	GAC	CAA	ATC	AGG	CTG	CAG	TAC	GGA	ATC	720
Ala	Leu	Val	Val	Gly	Glu	Val	Asp	Gln	Ile	Arg	Leu	Gln	Tyr	Gly	Ile	
225					230					235					240	
TTT	CGT	ATT	CAT	CAG	GAA	GTC	GAG	CCG	GAA	AAA	GGC	TCT	GAA	AAC	GCA	768
Phe	Arg	Ile	His	Gln	Glu	Val	Glu	Pro	Glu	Lys	Gly	Ser	Glu	Asn	Ala	
				245					250					255		
GTT	ATA	ACC	GTT	CCC	GCA	GAC	CTT	TCA	GCA	GAG	GAG	CGA	GGA	CGG	ATA	816
Val	Ile	Thr	Val	Pro	Ala	Asp	Leu	Ser	Ala	Glu	Glu	Arg	Gly	Arg	Ile	
			260				265						270			
CAG	GAA	ACG	GCA	AAA	AAA	ATA	TAT	AAA	GCG	CTC	GGC	TGT	AGA	GGT	CTA	864
Gln	Glu	Thr	Ala	Lys	Lys	Ile	Tyr	Lys	Ala	Leu	Gly	Cys	Arg	Gly	Leu	
		275					280					285				
GCC	CGT	GTG	GAT	ATG	TTT	TTA	CAA	GAT	AAC	GGC	CGC	ATT	GTA	CTG	AAC	912
Ala	Arg	Val	Asp	Met	Phe	Leu	Gln	Asp	Asn	Gly	Arg	Ile	Val	Leu	Asn	
	290					295					300					
GAA	GTC	AAT	ACT	CTG	CCC	GGT	TTC	ACG	TCA	TAC	AGT	CGT	TAT	CCC	CGT	960
Glu	Val	Asn	Thr	Leu	Pro	Gly	Phe	Thr	Ser	Tyr	Ser	Arg	Tyr	Pro	Arg	
305					310					315					320	

ATG	ATG	GCC	GCT	GCA	GGT	ATT	GCA	CTT	CCC	GAA	CTG	ATT	GAC	CGC	TTG	1008
Met	Met	Ala	Ala	Ala	Gly	Ile	Ala	Leu	Pro	Glu	Leu	Ile	Asp	Arg	Leu	
			325						330					335		

ATC	GTA	TTA	GCG	TTA	AAG	GGG	1029
Ile	Val	Leu	Ala	Leu	Lys	Gly	
			340				

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 343 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asn	Arg	Ile	Lys	Val	Ala	Ile	Leu	Phe	Gly	Gly	Cys	Ser	Glu	Glu	1	5	10	15
His	Asp	Val	Ser	Val	Lys	Ser	Ala	Ile	Glu	Ile	Ala	Ala	Asn	Ile	Asn	20	25	30	
Lys	Glu	Lys	Tyr	Glu	Pro	Leu	Tyr	Ile	Gly	Ile	Thr	Lys	Ser	Gly	Val	35	40	45	
Trp	Lys	Met	Cys	Glu	Lys	Pro	Cys	Ala	Glu	Trp	Glu	Asn	Asp	Asn	Cys	50	55	60	
Tyr	Ser	Ala	Val	Leu	Ser	Pro	Asp	Lys	Lys	Met	His	Gly	Leu	Leu	Val	65	70	75	80
Lys	Lys	Asn	His	Glu	Tyr	Glu	Ile	Asn	His	Val	Asp	Val	Ala	Phe	Ser	85	90	95	
Ala	Leu	His	Gly	Lys	Ser	Gly	Glu	Asp	Gly	Ser	Ile	Gln	Gly	Leu	Phe	100	105	110	
Glu	Leu	Ser	Gly	Ile	Pro	Phe	Val	Gly	Cys	Asp	Ile	Gln	Ser	Ser	Ala	115	120	125	
Ile	Cys	Met	Asp	Lys	Ser	Leu	Thr	Tyr	Ile	Val	Ala	Lys	Asn	Ala	Gly	130	135	140	
Ile	Ala	Thr	Pro	Ala	Phe	Trp	Val	Ile	Asn	Lys	Asp	Asp	Arg	Pro	Val	145	150	155	160
Ala	Ala	Thr	Phe	Thr	Tyr	Pro	Val	Phe	Val	Lys	Pro	Ala	Arg	Ser	Gly	165	170	175	
Ser	Ser	Phe	Gly	Val	Lys	Lys	Val	Asn	Ser	Ala	Asp	Glu	Leu	Asp	Tyr	180	185	190	
Ala	Ile	Glu	Ser	Ala	Arg	Gln	Tyr	Asp	Ser	Lys	Ile	Leu	Ile	Glu	Gln	195	200	205	

Ala Val Ser Gly Cys Glu Val Gly Cys Ala Val Leu Gly Asn Ser Ala
210 215 220

Ala Leu Val Val Gly Glu Val Asp Gln Ile Arg Leu Gln Tyr Gly Ile
225 230 235 240

Phe Arg Ile His Gln Glu Val Glu Pro Glu Lys Gly Ser Glu Asn Ala
245 250 255

Val Ile Thr Val Pro Ala Asp Leu Ser Ala Glu Glu Arg Gly Arg Ile
260 265 270

Gln Glu Thr Ala Lys Lys Ile Tyr Lys Ala Leu Gly Cys Arg Gly Leu
275 280 285

Ala Arg Val Asp Met Phe Leu Gln Asp Asn Gly Arg Ile Val Leu Asn
290 295 300

Glu Val Asn Thr Leu Pro Gly Phe Thr Ser Tyr Ser Arg Tyr Pro Arg
305 310 315 320

Met Met Ala Ala Ala Gly Ile Ala Leu Pro Glu Leu Ile Asp Arg Leu
325 330 335

Ile Val Leu Ala Leu Lys Gly
340

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 606 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..606

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG	GAA	ATA	GGA	TTT	ACT	TTT	TTA	GAT	GAA	ATA	GTA	CAC	GGT	GTT	CGT	48
Met	Glu	Ile	Gly	Phe	Thr	Phe	Leu	Asp	Glu	Ile	Val	His	Gly	Val	Arg	
1				5				10						15		
TGG	GAC	GCT	AAA	TAT	GCC	ACT	TGG	GAT	AAT	TTC	ACC	GGA	AAA	CCG	GTT	96
Trp	Asp	Ala	Lys	Tyr	Ala	Thr	Trp	Asp	Asn	Phe	Thr	Gly	Lys	Pro	Val	
			20					25					30			
GAC	GGT	TAT	GAA	GTA	AAT	CGC	ATT	GTA	GGG	ACA	TAC	GAG	TTG	GCT	GAA	144
Asp	Gly	Tyr	Glu	Val	Asn	Arg	Ile	Val	Gly	Thr	Tyr	Glu	Leu	Ala	Glu	
			35				40						45			

TCG	CTT	TTG	AAG	GCA	AAA	GAA	CTG	GCT	GCT	ACC	CAA	GGG	TAC	GGA	TTG	192
Ser	Leu	Leu	Lys	Ala	Lys	Glu	Leu	Ala	Ala	Thr	Gln	Gly	Tyr	Gly	Leu	
	50					55					60					
CTT	CTA	TGG	GAC	GGT	TAC	CGT	CCT	AAG	CGT	GCT	GTA	AAC	TGT	TTT	ATG	240
Leu	Leu	Trp	Asp	Gly	Tyr	Arg	Pro	Lys	Arg	Ala	Val	Asn	Cys	Phe	Met	
65					70					75					80	
CAA	TGG	GCT	GCA	CAG	CCG	GAA	AAT	AAC	CTG	ACA	AAG	GAA	AGT	TAT	TAT	288
Gln	Trp	Ala	Ala	Gln	Pro	Glu	Asn	Asn	Leu	Thr	Lys	Glu	Ser	Tyr	Tyr	
				85					90					95		
CCC	AAT	ATT	GAC	CGA	ACT	GAG	ATG	ATT	TCA	AAA	GGA	TAC	GTG	GCT	TCA	336
Pro	Asn	Ile	Asp	Arg	Thr	Glu	Met	Ile	Ser	Lys	Gly	Tyr	Val	Ala	Ser	
			100					105					110			
AAA	TCA	AGC	CAT	AGC	CGC	GGC	AGT	GCC	ATT	GAT	CTT	ACG	CTT	TAT	CGA	384
Lys	Ser	Ser	His	Ser	Arg	Gly	Ser	Ala	Ile	Asp	Leu	Thr	Leu	Tyr	Arg	
		115					120					125				
TTA	GAC	ACG	GGT	GAG	CTT	GTA	CCA	ATG	GGG	AGC	CGA	TTT	GAT	TTT	ATG	432
Leu	Asp	Thr	Gly	Glu	Leu	Val	Pro	Met	Gly	Ser	Arg	Phe	Asp	Phe	Met	
	130					135					140					
GAT	GAA	CGC	TCT	CAT	CAT	GCG	GCA	AAT	GGA	ATA	TCA	TGC	AAT	GAA	GCG	480
Asp	Glu	Arg	Ser	His	His	Ala	Ala	Asn	Gly	Ile	Ser	Cys	Asn	Glu	Ala	
145					150					155					160	
CAA	AAT	CGC	AGA	CGT	TTG	CGC	TCC	ATC	ATG	GAA	AAC	AGT	GGG	TTT	GAA	528
Gln	Asn	Arg	Arg	Arg	Leu	Arg	Ser	Ile	Met	Glu	Asn	Ser	Gly	Phe	Glu	
				165					170					175		
GCA	TAT	AGC	CTC	GAA	TGG	TGG	CAC	TAT	GTA	TTA	AGA	GAC	GAA	CCA	TAC	576
Ala	Tyr	Ser	Leu	Glu	Trp	Trp	His	Tyr	Val	Leu	Arg	Asp	Glu	Pro	Tyr	
			180					185					190			
CCC	AAT	AGC	TAT	TTT	GAT	TTC	CCC	GTT	AAA							606
Pro	Asn	Ser	Tyr	Phe	Asp	Phe	Pro	Val	Lys							
		195					200									

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 202 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Glu	Ile	Gly	Phe	Thr	Phe	Leu	Asp	Glu	Ile	Val	His	Gly	Val	Arg	
1				5					10					15		
Trp	Asp	Ala	Lys	Tyr	Ala	Thr	Trp	Asp	Asn	Phe	Thr	Gly	Lys	Pro	Val	
		20						25					30			

Asp Gly Tyr Glu Val Asn Arg Ile Val Gly Thr Tyr Glu Leu Ala Glu
 35 40 45
 Ser Leu Leu Lys Ala Lys Glu Leu Ala Ala Thr Gln Gly Tyr Gly Leu
 50 55 60
 Leu Leu Trp Asp Gly Tyr Arg Pro Lys Arg Ala Val Asn Cys Phe Met
 65 70 75 80
 Gln Trp Ala Ala Gln Pro Glu Asn Asn Leu Thr Lys Glu Ser Tyr Tyr
 85 90 95
 Pro Asn Ile Asp Arg Thr Glu Met Ile Ser Lys Gly Tyr Val Ala Ser
 100 105 110
 Lys Ser Ser His Ser Arg Gly Ser Ala Ile Asp Leu Thr Leu Tyr Arg
 115 120 125
 Leu Asp Thr Gly Glu Leu Val Pro Met Gly Ser Arg Phe Asp Phe Met
 130 135 140
 Asp Glu Arg Ser His His Ala Ala Asn Gly Ile Ser Cys Asn Glu Ala
 145 150 155 160
 Gln Asn Arg Arg Arg Leu Arg Ser Ile Met Glu Asn Ser Gly Phe Glu
 165 170 175
 Ala Tyr Ser Leu Glu Trp Trp His Tyr Val Leu Arg Asp Glu Pro Tyr
 180 185 190
 Pro Asn Ser Tyr Phe Asp Phe Pro Val Lys
 195 200

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1347 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 215..1243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGCTCTTCC TTCAACGCAC TTCTGTACCA AGAGTTGTTG TCCATTTGAT CACTAACAAT	60
AGCTTCCCCT GCTTTCTTCA AGCCCTTTGT CATAAAATCG TTAGATTTTC ATCATAAAAA	120
TACGAGAAAG ACAACAGGAA GACCGCAAAT TTTCTTTTCT TTCCTAGGT AACTGAATG	180

TAACCTTAAA AGAAAAAAGG AAAGGAAGAA AATG ATG AAA AAA ATT GCC GTT	232
Met Lys Lys Ile Ala Val	
1 5	
TTA TTT GGA GGG AAT TCT CCA GAA TAC TCA GTG TCA CTA ACC TCA GCA	280
Leu Phe Gly Gly Asn Ser Pro Glu Tyr Ser Val Ser Leu Thr Ser Ala	
10 15 20	
GCA AGT GTG ATC CAA GCT ATT GAC CCG CTG AAA TAT GAA GTA ATG ACC	328
Ala Ser Val Ile Gln Ala Ile Asp Pro Leu Lys Tyr Glu Val Met Thr	
25 30 35	
ATT GGC ATC GCA CCA ACA ATG GAT TGG TAT TGG TAT CAA GGA AAC CTC	376
Ile Gly Ile Ala Pro Thr Met Asp Trp Tyr Trp Tyr Gln Gly Asn Leu	
40 45 50	
GCG AAT GTT CGC AAT GAT ACT TGG CTA GAA GAT CAC AAA AAC TGT CAC	424
Ala Asn Val Arg Asn Asp Thr Trp Leu Glu Asp His Lys Asn Cys His	
55 60 65 70	
CAG CTG ACT TTT TCT AGC CAA GGA TTT ATA TTA GGA GAA AAA CGA ATC	472
Gln Leu Thr Phe Ser Ser Gln Gly Phe Ile Leu Gly Glu Lys Arg Ile	
75 80 85	
GTC CCT GAT GTC CTC TTT CCA GTC TTG CAT GGG AAG TAT GGC GAG GAT	520
Val Pro Asp Val Leu Phe Pro Val Leu His Gly Lys Tyr Gly Glu Asp	
90 95 100	
GGC TGT ATC CAA GGA CTG CTT GAA CTA ATG AAC CTG CCT TAT GTT GGT	568
Gly Cys Ile Gln Gly Leu Leu Glu Leu Met Asn Leu Pro Tyr Val Gly	
105 110 115	
TGC CAT GTC GCT GCC TCC GCA TTA TGT ATG AAC AAA TGG CTC TTG CAT	616
Cys His Val Ala Ala Ser Ala Leu Cys Met Asn Lys Trp Leu Leu His	
120 125 130	
CAA CTT GCT GAT ACC ATG GGA ATC GCT AGT GCT CCC ACT TTG CTT TTA	664
Gln Leu Ala Asp Thr Met Gly Ile Ala Ser Ala Pro Thr Leu Leu Leu	
135 140 145 150	
TCC CGC TAT GAA AAC GAT CCT GCC ACA ATC GAT CGT TTT ATT CAA GAC	712
Ser Arg Tyr Glu Asn Asp Pro Ala Thr Ile Asp Arg Phe Ile Gln Asp	
155 160 165	
CAT GGA TTC CCG ATC TTT ATC AAG CCG AAT GAA GCC GGT TCT TCA AAA	760
His Gly Phe Pro Ile Phe Ile Lys Pro Asn Glu Ala Gly Ser Ser Lys	
170 175 180	
GGG ATC ACA AAA GTA ACT GAC AAA ACA GCG CTC CAA TCT GCA TTA ACG	808
Gly Ile Thr Lys Val Thr Asp Lys Thr Ala Leu Gln Ser Ala Leu Thr	
185 190 195	
ACT GCT TTT GCT TAC GGT TCT ACT GTG TTG ATC CAA AAG GCG ATA GCG	856
Thr Ala Phe Ala Tyr Gly Ser Thr Val Leu Ile Gln Lys Ala Ile Ala	
200 205 210	

GGT ATT GAA ATT GGC TGC GGC ATC TTA GGA AAT GAG CAA TTG ACG ATT	904
Gly Ile Glu Ile Gly Cys Gly Ile Leu Gly Asn Glu Gln Leu Thr Ile	
215 220 225 230	
GGT GCT TGT GAT GCG ATT TCT CTT GTC GAC GGT TTT TTT GAT TTT GAA	952
Gly Ala Cys Asp Ala Ile Ser Leu Val Asp Gly Phe Phe Asp Phe Glu	
235 240 245	
GAG AAA TAC CAA TTA ATC AGC GCC ACG ATC ACT GTC CCA GCA CCA TTG	1000
Glu Lys Tyr Gln Leu Ile Ser Ala Thr Ile Thr Val Pro Ala Pro Leu	
250 255 260	
CCT CTC GCG CTT GAA TCA CAG ATC AAG GAG CAG GCA CAG CTG CTT TAT	1048
Pro Leu Ala Leu Glu Ser Gln Ile Lys Glu Gln Ala Gln Leu Leu Tyr	
265 270 275	
CGA AAC TTG GGA TTG ACG GGT CTG GCT CGA ATC GAT TTT TTC GTC ACC	1096
Arg Asn Leu Gly Leu Thr Gly Leu Ala Arg Ile Asp Phe Phe Val Thr	
280 285 290	
AAT CAA GGA GCG ATT TAT TTA AAC GAA ATC AAC ACC ATG CCG GGA TTT	1144
Asn Gln Gly Ala Ile Tyr Leu Asn Glu Ile Asn Thr Met Pro Gly Phe	
295 300 305 310	
ACT GGG CAC TCC CGC TAC CCA GCT ATG ATG GCG GAA GTC GGG TTA TCC	1192
Thr Gly His Ser Arg Tyr Pro Ala Met Met Ala Glu Val Gly Leu Ser	
315 320 325	
TAC GAA ATA TTA GTA GAG CAA TTG ATT GCA CTG GCA GAG GAG GAC AAA	1240
Tyr Glu Ile Leu Val Glu Gln Leu Ile Ala Leu Ala Glu Glu Asp Lys	
330 335 340	
CGA TGAACACATT ACAATTGATC AATAAAAACC ATCCATTGAA AAAAAATCAA	1293
Arg	
GAGCCCCCGC ACTTAGTGCT AGCTCCTTTT AGCGATCACG ATGTTTACCT GCAG	1347

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 343 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Lys Ile Ala Val Leu Phe Gly Gly Asn Ser Pro Glu Tyr Ser
1 5 10 15
Val Ser Leu Thr Ser Ala Ala Ser Val Ile Gln Ala Ile Asp Pro Leu
20 25 30
Lys Tyr Glu Val Met Thr Ile Gly Ile Ala Pro Thr Met Asp Trp Tyr
35 40 45

.Trp Tyr Gln Gly Asn Leu Ala Asn Val Arg Asn Asp Thr Trp Leu Glu
 50 55 60
 Asp His Lys Asn Cys His Gln Leu Thr Phe Ser Ser Gln Gly Phe Ile
 65 70 75 80
 Leu Gly Glu Lys Arg Ile Val Pro Asp Val Leu Phe Pro Val Leu His
 85 90 95
 Gly Lys Tyr Gly Glu Asp Gly Cys Ile Gln Gly Leu Leu Glu Leu Met
 100 105 110
 Asn Leu Pro Tyr Val Gly Cys His Val Ala Ala Ser Ala Leu Cys Met
 115 120 125
 Asn Lys Trp Leu Leu His Gln Leu Ala Asp Thr Met Gly Ile Ala Ser
 130 135 140
 Ala Pro Thr Leu Leu Leu Ser Arg Tyr Glu Asn Asp Pro Ala Thr Ile
 145 150 155 160
 Asp Arg Phe Ile Gln Asp His Gly Phe Pro Ile Phe Ile Lys Pro Asn
 165 170 175
 Glu Ala Gly Ser Ser Lys Gly Ile Thr Lys Val Thr Asp Lys Thr Ala
 180 185 190
 Leu Gln Ser Ala Leu Thr Thr Ala Phe Ala Tyr Gly Ser Thr Val Leu
 195 200 205
 Ile Gln Lys Ala Ile Ala Gly Ile Glu Ile Gly Cys Gly Ile Leu Gly
 210 215 220
 Asn Glu Gln Leu Thr Ile Gly Ala Cys Asp Ala Ile Ser Leu Val Asp
 225 230 235 240
 Gly Phe Phe Asp Phe Glu Glu Lys Tyr Gln Leu Ile Ser Ala Thr Ile
 245 250 255
 Thr Val Pro Ala Pro Leu Pro Leu Ala Leu Glu Ser Gln Ile Lys Glu
 260 265 270
 Gln Ala Gln Leu Leu Tyr Arg Asn Leu Gly Leu Thr Gly Leu Ala Arg
 275 280 285
 Ile Asp Phe Phe Val Thr Asn Gln Gly Ala Ile Tyr Leu Asn Glu Ile
 290 295 300
 Asn Thr Met Pro Gly Phe Thr Gly His Ser Arg Tyr Pro Ala Met Met
 305 310 315 320
 Ala Glu Val Gly Leu Ser Tyr Glu Ile Leu Val Glu Gln Leu Ile Ala
 325 330 335
 Leu Ala Glu Glu Asp Lys Arg
 340

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGNGARGAYG GNWSNHTNCA RGGN

24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AAYACNHTNC CNGGNTTYAC

20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 693 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..693

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATG AGC GAT AAA ATA CTT ATT GTG GAT GAT GAA CAT GAA ATT GCC GAT
Met Ser Asp Lys Ile Leu Ile Val Asp Asp Glu His Glu Ile Ala Asp
1 5 10 15

48

TTG GTT GAA TTA TAC TTA AAA AAC GAG AAT TAT ACG GTT TTC AAA TAC
Leu Val Glu Leu Tyr Leu Lys Asn Glu Asn Tyr Thr Val Phe Lys Tyr
20 25 30

96

[illegible]

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1152 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1152

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTG GTT ATA AAA TTG AAA AAT AAA AAA AAC GAC TAT TCC AAA CTA GAA	48
Leu Val Ile Lys Leu Lys Asn Lys Lys Asn Asp Tyr Ser Lys Leu Glu	
1 5 10 15	
CGA AAA CTT TAC ATG TAT ATC GTT GCA ATT GTT GTG GTA GCA ATT GTA	96
Arg Lys Leu Tyr Met Tyr Ile Val Ala Ile Val Val Val Ala Ile Val	
20 25 30	
TTC GTG TTG TAT ATT CGT TCA ATG ATC CGA GGG AAA CTT GGG GAT TGG	144
Phe Val Leu Tyr Ile Arg Ser Met Ile Arg Gly Lys Leu Gly Asp Trp	
35 40 45	
ATC TTA AGT ATT TTG GAA AAC AAA TAT GAC TTA AAT CAC CTG GAC GCG	192
Ile Leu Ser Ile Leu Glu Asn Lys Tyr Asp Leu Asn His Leu Asp Ala	
50 55 60	
ATG AAA TTA TAT CAA TAT TCC ATA CGG AAC AAT ATA GAT ATC TTT ATT	240
Met Lys Leu Tyr Gln Tyr Ser Ile Arg Asn Asn Ile Asp Ile Phe Ile	
65 70 75 80	
TAT GTG GCG ATT GTC ATT AGT ATT CTT ATT CTA TGT CGC GTC ATG CTT	288
Tyr Val Ala Ile Val Ile Ser Ile Leu Ile Leu Cys Arg Val Met Leu	
85 90 95	
TCA AAA TTC GCA AAA TAC TTT GAC GAG ATA AAT ACC GGC ATT GAT GTA	336
Ser Lys Phe Ala Lys Tyr Phe Asp Glu Ile Asn Thr Gly Ile Asp Val	
100 105 110	
CTT ATT CAG AAC GAA GAT AAA CAA ATT GAG CTT TCT GCG GAA ATG GAT	384
Leu Ile Gln Asn Glu Asp Lys Gln Ile Glu Leu Ser Ala Glu Met Asp	
115 120 125	
GTT ATG GAA CAA AAG CTC AAC ACA TTA AAA CGG ACT CTG GAA AAG CGA	432
Val Met Glu Gln Lys Leu Asn Thr Leu Lys Arg Thr Leu Glu Lys Arg	
130 135 140	
GAG CAG GAT GCA AAG CTG GCC GAA CAA AGA AAA AAT GAC GTT GTT ATG	480
Glu Gln Asp Ala Lys Leu Ala Glu Gln Arg Lys Asn Asp Val Val Met	
145 150 155 160	

TAC	TTG	GCG	CAC	GAT	ATT	AAA	ACG	CCC	CTT	ACA	TCC	ATT	ATC	GGT	TAT	528
Tyr	Leu	Ala	His	Asp	Ile	Lys	Thr	Pro	Leu	Thr	Ser	Ile	Ile	Gly	Tyr	
				165					170					175		
TTG	AGC	CTG	CTT	GAC	GAG	GCT	CCA	GAC	ATG	CCG	GTA	GAT	CAA	AAG	GCA	576
Leu	Ser	Leu	Leu	Asp	Glu	Ala	Pro	Asp	Met	Pro	Val	Asp	Gln	Lys	Ala	
			180					185					190			
AAG	TAT	GTG	CAT	ATC	ACG	TTG	GAC	AAA	GCG	TAT	CGA	CTC	GAA	CAG	CTA	624
Lys	Tyr	Val	His	Ile	Thr	Leu	Asp	Lys	Ala	Tyr	Arg	Leu	Glu	Gln	Leu	
		195					200					205				
ATC	GAC	GAG	TTT	TTT	GAG	ATT	ACA	CGG	TAT	AAC	CTA	CAA	ACG	ATA	ACG	672
Ile	Asp	Glu	Phe	Phe	Glu	Ile	Thr	Arg	Tyr	Asn	Leu	Gln	Thr	Ile	Thr	
	210					215					220					
CTA	ACA	AAA	ACG	CAC	ATA	GAC	CTA	TAC	TAT	ATG	CTG	GTG	CAG	ATG	ACC	720
Leu	Thr	Lys	Thr	His	Ile	Asp	Leu	Tyr	Tyr	Met	Leu	Val	Gln	Met	Thr	
					230					235					240	
GAT	GAA	TTT	TAT	CCT	CAG	CTT	TCC	GCA	CAT	GGA	AAA	CAG	GCG	GTT	ATT	768
Asp	Glu	Phe	Tyr	Pro	Gln	Leu	Ser	Ala	His	Gly	Lys	Gln	Ala	Val	Ile	
				245					250					255		
CAC	GCC	CCC	GAG	GAT	CTG	ACC	GTG	TCC	GGC	GAC	CCT	GAT	AAA	CTC	GCG	816
His	Ala	Pro	Glu	Asp	Leu	Thr	Val	Ser	Gly	Asp	Pro	Asp	Lys	Leu	Ala	
			260					265					270			
AGA	GTC	TTT	AAC	AAC	ATT	TTG	AAA	AAC	GCC	GCT	GCA	TAC	AGT	GAG	GAT	864
Arg	Val	Phe	Asn	Asn	Ile	Leu	Lys	Asn	Ala	Ala	Ala	Tyr	Ser	Glu	Asp	
		275					280					285				
AAC	AGC	ATC	ATT	GAC	ATT	ACC	GCG	GGC	CTC	TCC	GGG	GAT	GTG	GTG	TCA	912
Asn	Ser	Ile	Ile	Asp	Ile	Thr	Ala	Gly	Leu	Ser	Gly	Asp	Val	Val	Ser	
	290					295					300					
ATC	GAA	TTC	AAG	AAC	ACT	GGA	AGC	ATC	CCA	AAA	GAT	AAG	CTA	GCT	GCC	960
Ile	Glu	Phe	Lys	Asn	Thr	Gly	Ser	Ile	Pro	Lys	Asp	Lys	Leu	Ala	Ala	
	305				310					315				320		
ATA	TTT	GAA	AAG	TTC	TAT	AGG	CTG	GAC	AAT	GCT	CGT	TCT	TCC	GAT	ACG	1008
Ile	Phe	Glu	Lys	Phe	Tyr	Arg	Leu	Asp	Asn	Ala	Arg	Ser	Ser	Asp	Thr	
				325					330					335		
GGT	GGC	GCG	GGA	CTT	GGA	TTG	GCG	ATT	GCA	AAA	GAA	ATT	ATT	GTT	CAG	1056
Gly	Gly	Ala	Gly	Leu	Gly	Leu	Ala	Ile	Ala	Lys	Glu	Ile	Ile	Val	Gln	
			340					345					350			
CAT	GGA	GGG	CAG	ATT	TAC	GCG	GAA	AGC	AAT	GAT	AAC	TAT	ACG	ACG	TTT	1104
His	Gly	Gly	Gln	Ile	Tyr	Ala	Glu	Ser	Asn	Asp	Asn	Tyr	Thr	Thr	Phe	
		355					360					365				
AGG	GTA	GAG	CTT	CCA	GCG	ATG	CCA	GAC	TTG	GTT	GAT	AAA	AGG	AGG	TCC	1152
Arg	Val	Glu	Leu	Pro	Ala	Met	Pro	Asp	Leu	Val	Asp	Lys	Arg	Arg	Ser	
	370					375					380					

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Leu Val Ile Lys Leu Lys Asn Lys Lys Asn Asp Tyr Ser Lys Leu Glu
 1           5           10           15
Arg Lys Leu Tyr Met Tyr Ile Val Ala Ile Val Val Val Ala Ile Val
 20           25           30
Phe Val Leu Tyr Ile Arg Ser Met Ile Arg Gly Lys Leu Gly Asp Trp
 35           40           45
Ile Leu Ser Ile Leu Glu Asn Lys Tyr Asp Leu Asn His Leu Asp Ala
 50           55           60
Met Lys Leu Tyr Gln Tyr Ser Ile Arg Asn Asn Ile Asp Ile Phe Ile
 65           70           75           80
Tyr Val Ala Ile Val Ile Ser Ile Leu Ile Leu Cys Arg Val Met Leu
 85           90           95
Ser Lys Phe Ala Lys Tyr Phe Asp Glu Ile Asn Thr Gly Ile Asp Val
100           105           110
Leu Ile Gln Asn Glu Asp Lys Gln Ile Glu Leu Ser Ala Glu Met Asp
115           120           125
Val Met Glu Gln Lys Leu Asn Thr Leu Lys Arg Thr Leu Glu Lys Arg
130           135           140
Glu Gln Asp Ala Lys Leu Ala Glu Gln Arg Lys Asn Asp Val Val Met
145           150           155           160
Tyr Leu Ala His Asp Ile Lys Thr Pro Leu Thr Ser Ile Ile Gly Tyr
165           170           175
Leu Ser Leu Leu Asp Glu Ala Pro Asp Met Pro Val Asp Gln Lys Ala
180           185           190
Lys Tyr Val His Ile Thr Leu Asp Lys Ala Tyr Arg Leu Glu Gln Leu
195           200           205
Ile Asp Glu Phe Phe Glu Ile Thr Arg Tyr Asn Leu Gln Thr Ile Thr
210           215           220
Leu Thr Lys Thr His Ile Asp Leu Tyr Tyr Met Leu Val Gln Met Thr
225           230           235           240
Asp Glu Phe Tyr Pro Gln Leu Ser Ala His Gly Lys Gln Ala Val Ile
245           250           255
```

-His Ala Pro Glu Asp Leu Thr Val Ser Gly Asp Pro Asp Lys Leu Ala
 260 265 270
 Arg Val Phe Asn Asn Ile Leu Lys Asn Ala Ala Ala Tyr Ser Glu Asp
 275 280 285
 Asn Ser Ile Ile Asp Ile Thr Ala Gly Leu Ser Gly Asp Val Val Ser
 290 295 300
 Ile Glu Phe Lys Asn Thr Gly Ser Ile Pro Lys Asp Lys Leu Ala Ala
 305 310 315 320
 Ile Phe Glu Lys Phe Tyr Arg Leu Asp Asn Ala Arg Ser Ser Asp Thr
 325 330 335
 Gly Gly Ala Gly Leu Gly Leu Ala Ile Ala Lys Glu Ile Ile Val Gln
 340 345 350
 His Gly Gly Gln Ile Tyr Ala Glu Ser Asn Asp Asn Tyr Thr Thr Phe
 355 360 365
 Arg Val Glu Leu Pro Ala Met Pro Asp Leu Val Asp Lys Arg Arg Ser
 370 375 380

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7225 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGCTTTTCT TTTTGCTCAT TTGTTAGAGA TTTACTAACC GTATTAAATA GCTTCTTTTC	60
AGCCATTGCC CTTGCTTCCC ACACCATTCT TTCAAGTGTA GTGATAGCAG GCAGTATAAT	120
TTTGTTTTTT CTTAGAAAAT CTATGCATTC ATGCAGTAGA TGAATGGCAT CACCATTTTC	180
CAAAGCTAAT TGATGAAGGT ACTTAAATGT CATTCGATAT TCACTCAGGG TAAAAGTTAC	240
AAAGTCGTAT TCACTTCGAA TTTCTTTCAA ATGATCCCAA AGTGTATTTT CCCTTTGAGG	300
ATAATGATCA AGCGAGGATG GACTAACACC AATCTGTTTC GATATATATT GTATGACCGA	360
ATCTGGGATG CTTTTGATAT GAGTGTATGG CCAACCGGGA TACCGAAGAA CAGCTAATTG	420
AACAGCAAAT CCTAAACGGT TTTCTTCCCT CCTTCGCTTA TTA ACTATTT CTAAATCCCG	480
TTTGGA AAAA GTGAAGTAGG TCCCCAGTAT CCATTCATCT TCAGGGATTT GCATAAAAGC	540
CTGTCTCTGT TCCGGTGTA GCAATTCTCT ACCTCTCGCA ATTTTCATTC AGTATCATTC	600

CATTTCTGTA	TTTTCAATTT	ATTAGTTCAA	TTATATATCA	ATAGAGTGTA	CTCTATTGAT	660
ACAAATGTAG	TAGACTGATA	AAATCATAGT	TAAGAGCGTC	TCATAAGACT	TGTCTCAAAA	720
ATGAGGTGAT	ATTTTGCGGA	AAATCGGTTA	TATTCGTGTC	AGTTCGACTA	ACCAGAATCC	780
TTCAAGACAA	TTTCAGCAGT	TGAACGAGAT	CGGAATGGAT	ATTATATAAA	GAGAAAGTTT	840
CAGGAGCAAC	AAAGGATCGC	GAGCAACTTC	AAAAAGTGTT	AGACGATTTA	CAGGAAGATG	900
ACATCATTTA	TGTTACAGAC	TTAACTCGAA	TCACTCGTAG	TACACAAGAT	CTATTTGAAT	960
TAATCGATAA	CATACGAGAT	AAAAAGGCAA	GTTTAAATC	ACTAAAAGAT	ACATGGCTTG	1020
ATTTATCAGA	AGATAATCCA	TACAGCCAAT	TCTTAATTAC	TGTAATGGCT	GGTGTTAACC	1080
AATTAGAGCG	AGATCTTATT	CGGATGAGAC	AACGTGAAGG	GATTGAATTG	GCTAAGAAAG	1140
AAGGAAAGTT	TAAAGGTCGA	TTAAAGAAGT	ATCATAAAAA	TCACGCAGGA	ATGAATTATG	1200
CGGAAAGCTA	TATAAAGAAG	GAAATATGAC	TGTAAATCAA	ATTTGTGAAA	TTACTAATGT	1260
ATCTAGGGCT	TCATTATACA	GGAAATTATC	AGAAGTGAAT	AATTAGCCAT	TCTGTATTCC	1320
GCTAATGGGC	AATATTTTTA	AAGAAGAAAA	GGAAACTATA	AAATATTAAC	AGCCTCCTAG	1380
CGATGCCGAA	AAGCCCTTTG	ATAAAAAAAG	AATCATCATC	TTAAGAAATT	CTTAGTCATT	1440
TATTATGTAA	ATGCTTATAA	ATTCGGCCCT	ATAATCTGAT	AAATTATTAA	GGGCAAACCTT	1500
ATGTGAAAGG	GTGATAACTA	TGAGCGATAA	AATACTTATT	GTGGATGATG	AACATGAAAT	1560
TGCCGATTTG	GTTGAATTAT	ACTTAAAAAA	CGAGAATTAT	ACGGTTTTCA	AATACTATAC	1620
CGCCAAAGAA	GCATTGGAAT	GTATAGACAA	GTCTGAGATT	GACCTTGCCA	TATTGGACAT	1680
CATGCTTCCC	GGCACAAGCG	GCCTTACTAT	CTGTCAAAAA	ATAAGGGACA	AGCACACCTA	1740
TCCGATTATC	ATGCTGACCG	GGAAAGATAC	AGAGGTAGAT	AAAATTACAG	GGTTAACAAT	1800
CGGCGCGGAT	GATTATATAA	CGAAGCCCTT	TCGCCCACTG	GAGTTAATTG	CTCGGGTAAA	1860
GGCCCAGTTG	CGCCGATACA	AAAAATTCAG	TGGAGTAAAG	GAGCAGAACG	AAAATGTTAT	1920
CGTCCACTCC	GGCCTTGTCA	TTAATGTTAA	CACCCATGAG	TGTTATCTGA	ACGAGAAGCA	1980
GTTATCCCTT	ACTCCCACCG	AGTTTTCAAT	ACTGCGAATC	CTCTGTGAAA	ACAAGGGGAA	2040
TGTGGTTAGC	TCCGAGCTGC	TATTTTCATGA	GATATGGGGC	GACGAATATT	TCAGCAAGAG	2100
CAACAACACC	ATCACCGTGC	ATATCCGGCA	TTTGCGCGAA	AAAATGAACG	ACACCATTGA	2160
TAATCCGAAA	TATATAAAAA	CGGTATGGGG	GGTTGGTTAT	AAAATTGAAA	AATAAAAAAA	2220
ACGACTATTC	CAAAC TAGAA	CGAAAAC TTT	ACATGTATAT	CGTTGCAATT	GTTGTGGTAG	2280
CAATTGTATT	CGTGTTGTAT	ATTCGTTCAA	TGATCCGAGG	GAAACTTGGG	GATTGGATCT	2340

TAAGTATTTT	GGAAAACAAA	TATGACTTAA	ATCACCTGGA	CGCGATGAAA	TTATATCAAT	2400
ATTCCATACG	GAACAATATA	GATATCTTTA	TTTATGTGGC	GATTGTCATT	AGTATTCTTA	2460
TTCTATGTCG	CGTCATGCTT	TCAAAATTCTG	CAAATACTT	TGACGAGATA	AATACCGGCA	2520
TTGATGTACT	TATTCAGAAC	GAAGATAAAC	AAATTGAGCT	TTCTGCGGAA	ATGGATGTTA	2580
TGGAACAAAA	GCTCAACACA	TTAAAACGGA	CTCTGGAAAA	GCGAGAGCAG	GATGCAAAGC	2640
TGGCCGAACA	AAGAAAAAAT	GACGTTGTTA	TGTACTTGGC	GCACGATATT	AAAACGCCCC	2700
TTACATCCAT	TATCGGTTAT	TTGAGCCTGC	TTGACGAGGC	TCCAGACATG	CCGGTAGATC	2760
AAAAGGCAAA	GTATGTGCAT	ATCACGTTGG	ACAAAGCGTA	TCGACTCGAA	CAGCTAATCG	2820
ACGAGTTTTT	TGAGATTACA	CGGTATAACC	TACAAACGAT	AACGCTAACA	AAAACGCACA	2880
TAGACCTATA	CTATATGCTG	GTGCAGATGA	CCGATGAATT	TTATCCTCAG	CTTCCGCAC	2940
ATGGAAAACA	GGCGGTTATT	CACGCCCCCG	AGGATCTGAC	CGTGTCCGGC	GACCCTGATA	3000
AACTCGCGAG	AGTCTTTAAC	AACATTTTGA	AAAACGCCGC	TGCATACAGT	GAGGATAACA	3060
GCATCATTGA	CATTACCGCG	GGCCTCTCCG	GGGATGTGGT	GTCAATCGAA	TTCAAGAACA	3120
CTGGAAGCAT	CCCAAAGAT	AAGCTAGCTG	CCATATTTGA	AAAGTTCTAT	AGGCTGGACA	3180
ATTCTCGTTC	TTCCGATACG	GGTGGCGCGG	GACTTGGAAT	GGCGATTGCA	AAAGAAATTA	3240
TTGTTCAGCA	TGGAGGGCAG	ATTTACGCGG	AAAGCTATGA	TAACTATACG	ACGTTTAGGG	3300
TAGAGCTTCC	AGCGATGCCA	GACTTGGTTG	ATAAAAGGAG	GTCCTAAGAG	ATGTATATAA	3360
TTTTTTAGGA	AAATCTCAAG	GTTATCTTTA	CTTTTTCTTA	GGAAATTAAC	AATTTAATAT	3420
TAAGAAACGG	CTCGTTCTTA	CACGGTAGAC	TTAATACCGT	AAGAACGAGC	CGTTTTTCGTT	3480
CTTCAGAGAA	AGATTTGACA	AGATTACCAT	TGGCATCCCC	GTTTTATTTG	GTGCCTTTCA	3540
CAGAAAGGGT	TGGTCTTAAT	TATGAATAAC	ATCGGCATTA	CTGTTTATGG	ATGTGAGCAG	3600
GATGAGGCAG	ATGCATTCCA	TGCTCTTTTCG	CCTCGCTTTG	GCGTTATGGC	AACGATAATT	3660
AACGCCAACG	TGTCGGAATC	CAACGCCAAA	TCCGCGCCTT	TCAATCAATG	TATCAGTGTG	3720
GGACATAAAT	CAGAGATTTT	CGCCTCTATT	CTTCTTGCGC	TGAAGAGAGC	CGGTGTGAAA	3780
TATATTTCTA	CCCGAAGCAT	CGGCTGCAAT	CATATAGATA	CAACTGCTGC	TAAGAGAATG	3840
GGCATCACTG	TCGACAATGT	GGCGTACTCG	CCGGATAGCG	TTGCCGATTA	TACTATGATG	3900
CTAATTCTTA	TGGCAGTACG	CAACGTAAAA	TCGATTGTGC	GCTCTGTGGA	AAAACATGAT	3960
TTCAGGTTGG	ACAGCGACCG	TGGCAAGGTA	CTCAGCGACA	TGACAGTTGG	TGTGGTGGGA	4020
ACGGGCCAGA	TAGGCAAAGC	GGTTATTGAG	CGGCTGCGAG	GATTTGGATG	TAAAGTGTG	4080

GCTTATAGTC	GCAGCCGAAG	TATAGAGGTA	AACTATGTAC	CGTTTGATGA	GTTGATGCAA	4140
AATAGCGATA	TCGTTACGCT	TCATGTGCCG	CTCAATACGG	ATACGCACTA	TATTATCAGC	4200
CACGAACAAA	TACAGAGAAT	GAAGCAAGGA	GCATTTCTTA	TCAATACTGG	GCGCGGTCCA	4260
CTTGTAAGATA	CCTATGAGTT	GGTTAAAGCA	TTAGAAAACG	GGAAACTGGG	CGGTGCCGCA	4320
TTGGATGTAT	TGGAAGGAGA	GGAAGAGTTT	TTCTACTCTG	ATTGCACCCA	AAAACCAATT	4380
GATAATCAAT	TTTTACTTAA	ACTTCAAAGA	ATGCCTAACG	TGATAATCAC	ACCGCATACG	4440
GCCTATTATA	CCGAGCAAGC	GTTGCGTGAT	ACCGTTGAAA	AAACCATTAA	AAACTGTTTG	4500
GATTTTGAAA	GGAGACAGGA	GCATGAATAG	AATAAAAGTT	GCAATACTGT	TTGGGGGTTG	4560
CTCAGAGGAG	CATGACGTAT	CGGTAAAATC	TGCAATAGAG	ATAGCCGCTA	ACATTAATAA	4620
AGAAAAATAC	GAGCCGTTAT	ACATTGGAAT	TACGAAATCT	GGTGTATGGA	AAATGTGCGA	4680
AAAACCTTGC	GCGGAATGGG	AAAACGACAA	TTGCTATTCA	GCTGTACTCT	CGCCGGATAA	4740
AAAAATGCAC	GGATTACTTG	TTAAAAAGAA	CCATGAATAT	GAAATCAACC	ATGTTGATGT	4800
AGCATTTTCA	GCTTTGCATG	GCAAGTCAGG	TGAAGATGGA	TCCATACAAG	GTCTGTTTGA	4860
ATTGTCCGGT	ATCCCTTTTG	TAGGCTGCGA	TATTCAAAGC	TCAGCAATTT	GTATGGACAA	4920
ATCGTTGACA	TACATCGTTG	CGAAAAATGC	TGGGATAGCT	ACTCCCGCCT	TTTGGGTTAT	4980
TAATAAAGAT	GATAGGCCGG	TGGCAGCTAC	GTTTACCTAT	CCTGTTTTTG	TTAAGCCGGC	5040
GCGTTCAGGC	TCATCCTTCG	GTGTGAAAAA	AGTCAATAGC	GCGGACGAAT	TGGACTIONGC	5100
AATTGAATCG	GCAAGACAAT	ATGACAGCAA	AATCTTAATT	GAGCAGGCTG	TTTCGGGCTG	5160
TGAGGTCGGT	TGTGCGGTAT	TGGGAAACAG	TGCCGCGTTA	GTTGTTGGCG	AGGTGGACCA	5220
AATCAGGCTG	CAGTACGGAA	TCTTTCGTAT	TCATCAGGAA	GTCGAGCCGG	AAAAAGGCTC	5280
TGAAAACGCA	GTTATAACCG	TTCCCGCAGA	CCTTTCAGCA	GAGGAGCGAG	GACGGATACA	5340
GGAAACGGCA	AAAAAAATAT	ATAAAGCGCT	CGGCTGTAGA	GGTCTAGCCC	GTGTGGATAT	5400
GTTTTTACAA	GATAACGGCC	GCATTGTACT	GAACGAAGTC	AATACTCTGC	CCGGTTTCAC	5460
GTCATACAGT	CGTTATCCCC	GTATGATGGC	CGCTGCAGGT	ATTGCACTTC	CCGAACTGAT	5520
TGACCGCTTG	ATCGTATTAG	CGTTAAAGGG	GTGATAAGCA	TGGAAATAGG	ATTTACTTTT	5580
TTAGATGAAA	TAGTACACGG	TGTTTCGTTG	GACGCTAAAT	ATGCCACTTG	GGATAATTTT	5640
ACCGGAAAAC	CGGTTGACGG	TTATGAAGTA	AATCGCATTG	TAGGGACATA	CGAGTTGGCT	5700
GAATCGCTTT	TGAAGGCAAA	AGAACTGGCT	GCTACCCAAG	GGTACGGATT	GCTTCTATGG	5760
GACGGTTACC	GTCCTAAGCG	TGCTGTAAAC	TGTTTTATGC	AATGGGCTGC	ACAGCCGGAA	5820

.AATAACCTGA	CAAAGGAAAG	TTATTATCCC	AATATTGACC	GAAGTGAAGAT	GATTTCAAAA	5880
GGATACGTGG	CTTCAAAATC	AAGCCATAGC	CGCGGCAGTG	CCATTGATCT	TACGCTTTAT	5940
CGATTAGACA	CGGGTGAGCT	TGTACCAATG	GGGAGCCGAT	TTGATTTTAT	GGATGAACGC	6000
TCTCATCATG	CGGCAAATGG	AATATCATGC	AATGAAGCGC	AAAATCGCAG	ACGTTTGCGC	6060
TCCATCATGG	AAAACAGTGG	GTTTGAAGCA	TATAGCCTCG	AATGGTGGCA	CTATGTATTA	6120
AGAGACGAAC	CATACCCCAA	TAGCTATTTT	GATTTCCCCG	TTAAATAAAC	TTTTAACCGT	6180
TGCACGGACA	AACTATATAA	GCTAACTCTT	TCGGCAGGAA	ACCCGACGTA	TGTAAGTGGT	6240
TCTTAGGGAA	TTTATATATA	GTAGATAGTA	TTGAAGATGT	AAGGCAGAGC	GATATTGCGG	6300
TCATTATCTG	CGTGCGCTGC	GGCAAGATAG	CCTGATAATA	AGACTGATCG	CATAGAGGGG	6360
TGGTATTTCA	CACCGCCCAT	TGTCAACAGG	CAGTTCAGCC	TCGTTAAATT	CAGCATGGGT	6420
ATCACTTATG	AAAATTCATC	TACATTGGTG	ATAATAGTAA	ATCCAGTAGG	GCGAAATAAT	6480
TGACTGTAAT	TTACGGGGCA	AAACGGCACA	ATCTCAAACG	AGATTGTGCC	GTTTAAGGGG	6540
AAGATTCTAG	AAATATTTCA	TACTTCCAAC	TATATAGTTA	AGGAGGAGAC	TGAAAATGAA	6600
GAAGTTGTTT	TTTTTATTGT	TATTGTTATT	CTTAATATAC	TTAGGTATATG	ACTACGTTAA	6660
TGAAGCACTG	TTTTCTCAGG	AAAAAGTCGA	ATTTCAAAAT	TATGATCAAA	ATCCCAAAGA	6720
ACATTTAGAA	AATAGTGGGA	CTTCTGAAAA	TACCCAAGAG	AAAACAATTA	CAGAAGAACA	6780
GGTTTATCAA	GGAAATCTGC	TATTAATCAA	TAGTAAATAT	CCTGTTCGCC	AAGAAGTGTG	6840
AAGTCAGATA	TCGTGAATTT	ATCTAAACAT	GACGAATTAA	TAAATGGATA	CGGGTTGCTT	6900
GATAGTAATA	TTTATATGTC	AAAAGAAATA	GCACAAAAAT	TTTCAGAGAT	GGTCAATGAT	6960
GCTGTAAAGG	GTGGCGTTAG	TCATTTTATT	ATTAATAGTG	GCTATCGAGA	CTTTGATGAG	7020
CAAAGTGTGC	TTTACCAAGA	AATGGGGGCT	GAGTATGCCT	TACCAGCAGG	TTATAGTGAG	7080
CATAATTCAG	GTTTATCACT	AGATGTAGGA	TCAAGCTTGA	CGAAAATGGA	ACGAGCCCCT	7140
GAAGGAAAGT	GGATAGAAGA	AAATGCTTGG	AAATACGGGT	TCATTTTACG	TTATCCAGAG	7200
GACAAAACAG	AGTTAACAGG	AATTC				7225

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10851 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGGTAGCGT CAGGAAAATG CGGATTTACA ACGCTAAGCC TATTTTCCTG ACGAATCCCT	60
CGTTTTTAAC AACGTTAAGA AAGTTTTAGT GGTCTTAAAG AATTTAATGA GACTACTTTC	120
TCTGAGTTAA AATGGTATTC TCCTAGTAAA TTAATATGTT CCCAACCTAA GGGCGACATA	180
TGGTGTAACA AATCTTCATT AAAGCTACCT GTCCGTTTTT TATATTCAAC TGCTGTTGTT	240
AGGTGGAGAG TATTCCAAAT ACTTATAGCA TTGATAATTA TGTTTAAAGC ACTGGCTCTT	300
TGCAATTGAT GCTGTATGGT GCGTTCTCTA AGCTCACCTT GTTTTCCGAA GAAAATAGCT	360
CTTGCCAATC CATTCAATGGC TTCTCCTTTA TTCAATCCTC TTTGTATTTT TCTTCTTAAT	420
GATTCATCCG ATATATAATT CAAAATAAAG ATCGTTTTTT CTATTCGGCC CATCTCACGT	480
AAGGCTGTAG CTAAGCTGTT TTGTCTTGAA TAGGAACCTA GCTTCCCCAT AATAAGGGAT	540
GCTGAAACTG TTCCCTCCCT TATAGAATGA GCTAATCGCA AAACATCCTC ATAATTTTCT	600
TTAATGACCT TTGTATTTAT TTGTCCACGT AAAATGGCTT CTAGTTTGGG ATACTCACTT	660
GCTTTATCTA TCGTAAATAA TTTTGAGTCC GATAAATCCC TTATTCTTGG GGCAAATTTA	720
AATCCTAATA AATGAGTCAG TCCGAATATT TGGTCAGTGT AACCGGCAGT GTCTGTATAA	780
TGTTCCCTCTA TGTTTAGATC CGTCTCATGA TGTAACAAAC CATCCAAAAC ATGAATCGCA	840
TCTCTTGAAT TAGTATGAAT AATCTTTGTG TAGTAAGAAG AGAATTGATC ACTTGTAAT	900
CGGTAGATGG TGGCTCCTTT TCCAGTTCCA TAATGTGGAT TTGCATCTGC ATGTAGTGAT	960
GAAACACCTA GCTGCATTCT CATAACATCT GACGAAGATG TTGTACCGTC GCCCAATAG	1020
AAAGGCAATT GTAATTTATG ATGAAAGTTT ACTAATATGG CTTGGGCTTT ATTCATGGCA	1080
TCTTCATACA TGCGCCATTG AGATACATTG GCTAGTTGCT TATATGTAAG TCCGGGTGTG	1140
GCTTCGGCCA TCTTGCTCAA GCCAATATTC ATTCCCATTCT CTAAGGGG AGCCATGATA	1200
ATGATTGTTT CTTCTTATC TGGTTTTCTGA TTATTGGAAG CATGAGTGAA TTGCTCATGA	1260
AATCCTGTTA TATGGGCCAC ATCCATGAGT AAATCAGTTA ATTTTATTCT TGGTAGCATC	1320
TGATAAAGGC TTGCACTAAA TTTTTTTGCT TCTTCTGGAA CATCTTTTTC TAAGCGTGCA	1380
AGTGATAGCT TTCCTTTTTC AAGAGAAACC CCATCTAACT TATTGGAATT GGCAGCTAAC	1440
CACTTTAACC TTTCATTAAA GCTGCTGGTT CTCTCCGTTA TATAATCTTC GAATGATAAA	1500
CTAACTGATA ATCTCGTATT CCCCTTCGAT TGATTCCATG TATCTTCCGA AAACAAATAT	1560
TCCTCAAAAT CCCTATATTG TCTGCTGCCA ACAATGGAAA CATCTCCTGC CCGAACATGC	1620

TCCCGAAGTT	CTGTAAAAAC	AGCCATTTCA	TAGTAATGAC	GATTAATTGT	TGTACCATCA	1680
TCCTCGTATA	AATGTCTTTT	CCATCGTTTT	GAAATAAAAT	CCACAGGTGA	GTCATCAGGC	1740
ACTTTTCGCT	TTCCAGATTC	GTTCAATCCT	CGGATAATCT	CAACAGCTTG	TAAAAGTGGC	1800
TCATTTGCCT	TTGTAGAATG	AAATTCCAAT	ACTCTTAATA	GCGTTGGCGT	ATATTTTCTT	1860
AGTGAATAAA	ACCGTTTTTG	CAGTAAGTCT	AAATAATCAT	AGTCGGCAGG	ACGTGCAAGT	1920
TCCTGAGCCT	CTTCTACTGA	AGAGACAAAG	GTATTCCATT	CAATAACCGA	TTCTAAAACC	1980
TTAAAAACGT	CTAATTTTTC	CTCTCTTGCT	TTAATTAATG	CTTGTCCGAT	GTTCGTAAAG	2040
TGTATAACTT	TCTCATTTAG	CTTTTTACCG	TTTTGTTTCT	GGATTTCCCTC	TTGAGCCTTA	2100
CGACCTTTTG	ATAACAAACT	AAGTATTTGC	CTATCATGAA	TTTCAAACGC	TTTATCCGTT	2160
AGCTCCTGAG	TAAGTTGTAA	TAAATAGATG	GTTAATATCG	AATAACGTTT	ATTTTCTTGA	2220
AAGTCACGGA	ATGCATACGG	CTCGTATCTT	GAGCCTAAGC	GAGACAGCTG	CAACAGGCGG	2280
TTACGGTGCA	AATGACTAAT	TTGCACTGTT	TCTAAATCCA	TTCTCGTAT	GTATTCGAGT	2340
CGTTCTATTA	TTTTTAGAAA	AGTTTCGGGT	GAAGGATGAC	CCGGTGGCTC	TTTTAACCAA	2400
CCCAATATCG	TTTTATTGGA	TTCCGATGGA	TGCTGCGAGG	TAATAATCCC	TTCAAGCTTT	2460
TCTTTTTGCT	CATTTGTTAG	AGATTTACTA	ACCGTATTAA	ATAGCTTCTT	TTCAGCCATT	2520
GCCCTTGCTT	CCCACACCAT	TCTTTCAAGT	GTAGTGATAG	CAGGCAGTAT	AATTTTGTTT	2580
TTTCTTAGAA	AATCTATGCA	TTCATGCAGT	AGATGAATGG	CATCACCATT	TTCCAAAGCT	2640
AATTGATGAA	GGTACTTAAA	TGTCATT CGA	TATTCACTCA	GGGTAAAAGT	TACAAAGTCG	2700
TATTCACTTC	GAATTTCTTT	CAAATGATCC	CAAAGTGTAT	TTTCCCTTTG	AGGATAATGA	2760
TCAAGCGAGG	ATGGACTAAC	ACCAATCTGT	TTCGATATAT	ATTGTATGAC	CGAATCTGGG	2820
ATGCTTTTGA	TATGAGTGTA	TGGCCAACCG	GGATACCGAA	GAACAGCTAA	TTGAACAGCA	2880
AATCCTAAAC	GGTTTTCTTC	CCTCCTTCGC	TTATTAACTA	TTTCTAAATC	CCGTTTGGA	2940
AAAGTGAAGT	AGGTCCCCAG	TATCCATTCA	TCTTCAGGGA	TTGCATAAA	AGCCTGTCTC	3000
TGTTCCGGTG	TAAGCAATTC	TCTACCTCTC	GCAATTTTCA	TTCAGTATCA	TTCCATTTCT	3060
GTATTTTCAA	TTTATTAGTT	CAATTATATA	TCAATAGAGT	GTA CTCTATT	GATACAAATG	3120
TAGTAGACTG	ATAAAATCAT	AGTTAAGAGC	GTCTCATAAG	ACTTGCTCTCA	AAAATGAGGT	3180
GATATTTTGC	GGAAAATCGG	TTATATTCGT	GTCAGTTCGA	CTAACCAGAA	TCCTTCAAGA	3240
CAATTTCAGC	AGTTGAACGA	GATCGGAATG	GATATTATAT	ATGAAGAGAA	AGTTTCAGGA	3300
GCAACAAAGG	ATCGCGAGCA	ACTTCAAAAA	GTGTTAGACG	ATTTACAGGA	AGATGACATC	3360

ATTTATGTTA	CAGACTTAAC	TCGAATCACT	CGTAGTACAC	AAGATCTATT	TGAATTAATC	3420
GATAACATAC	GAGATAAAAA	GGCAAGTTTA	AAATCACTAA	AAGATACATG	GCTTGATTTA	3480
TCAGAAGATA	ATCCATACAG	CCAATTCTTA	ATTACTGTAA	TGGCTGGTGT	TAACCAATTA	3540
GAGCGAGATC	TTATTCGGAT	GAGACAACGT	GAAGGGATTG	AATTGGCTAA	GAAAGAAGGA	3600
AAGTTTAAAG	GTCGATTAAA	GAAGTATCAT	AAAAATCACG	CAGGAATGAA	TTATGCGGTA	3660
AAGCTATATA	AAGAAGGAAA	TATGACTGTA	AATCAAATTT	GTGAAATTAC	TAATGTATCT	3720
AGGGCTTCAT	TATACAGGAA	ATTATCAGAA	GTGAATAATT	AGCCATTCTG	TATTCGCTA	3780
ATGGGCAATA	TTTTTAAAGA	AGAAAAGGAA	ACTATAAAAT	ATTAACAGCC	TCCTAGCGAT	3840
GCCGAAAAGC	CCTTTGATAA	AAAAAGAATC	ATCATCTTAA	GAAATTCTTA	GTCATTTATT	3900
ATGTAAATGC	TTATAAATTC	GGCCCTATAA	TCTGATAAAT	TATTAAGGGC	AAACTTATGT	3960
GAAAGGGTGA	TAACATATGAG	CGATAAAATA	CTTATTGTGG	ATGATGAACA	TGAAATTGCC	4020
GATTTGGTTG	AATTATACTT	AAAAAACGAG	AATTATACGG	TTTTCAAATA	CTATACCGCC	4080
AAAGAAGCAT	TGGAATGTAT	AGACAAGTCT	GAGATTGACC	TTGCCATATT	GGACATCATG	4140
CTTCCCGGCA	CAAGCGGCCT	TACTATCTGT	CAAAAAATAA	GGGACAAGCA	CACCTATCCG	4200
ATTATCATGC	TGACCGGGAA	AGATACAGAG	GTAGATAAAA	TTACAGGGTT	AACAATCGGC	4260
GCGGATGATT	ATATAACGAA	GCCCTTTCGC	CCACTGGAGT	TAATTGCTCG	GGTAAAGGCC	4320
CAGTTGCGCC	GATACAAAAA	ATTCAGTGGA	GTAAAGGAGC	AGAACGAAAA	TGTTATCGTC	4380
CACTCCGGCC	TTGTCATTAA	TGTTAACACC	CATGAGTGTT	ATCTGAACGA	GAAGCAGTTA	4440
TCCCTTACTC	CCACCGAGTT	TTCAATACTG	CGAATCCTCT	GTGAAAACAA	GGGGAATGTG	4500
GTTAGCTCCG	AGCTGCTATT	TCATGAGATA	TGGGGCGACG	AATATTTTCA	CAAGAGCAAC	4560
AACACCATCA	CCGTGCATAT	CCGGCATTTC	CGCGAAAAAA	TGAACGACAC	CATTGATAAT	4620
CCGAAATATA	TAAAAACGGT	ATGGGGGGTT	GGTTATAAAA	TTGAAAAATA	AAAAAACGAA	4680
CTATTCCAAA	CTAGAACGAA	AACTTTACAT	GTATATCGTT	GCAATTGTTG	TGGTAGCAAT	4740
TGTATTCGTG	TTGTATATTC	GTTCAATGAT	CCGAGGGAAA	CTTGGGGATT	GGATCTTAAG	4800
TATTTTGGAA	AACAAATATG	ACTTAAATCA	CCTGGACGCG	ATGAAATTAT	ATCAATATTC	4860
CATACGGAAC	AATATAGATA	TCTTTATTTA	TGTGGCGATT	GTCATTAGTA	TTCTTATTCT	4920
ATGTCGCGTC	ATGCTTTCAA	AATTCGCAAA	ATACTTTGAC	GAGATAAATA	CCGGCATTGA	4980
TGTACTTATT	CAGAACGAAG	ATAAACAAAT	TGAGCTTTCT	GCGGAAATGG	ATGTTATGGA	5040
ACAAAAGCTC	AACACATTAA	AACGGACTCT	GGAAAAGCGA	GAGCAGGATG	CAAAGCTGGC	5100

CGAACAAAGA	AAAAATGACG	TTGTTATGTA	CTTGGCGCAC	GATATTAAAA	CGCCCCTTAC	5160
ATCCATTATC	GGTTATTTGA	GCCTGCTTGA	CGAGGCTCCA	GACATGCCGG	TAGATCAAAA	5220
GGCAAAGTAT	GTGCATATCA	CGTTGGACAA	AGCGTATCGA	CTCGAACAGC	TAATCGACGA	5280
GTTTTTTGAG	ATTACACGGT	ATAACCTACA	AACGATAACG	CTAACAAAAA	CGCACATAGA	5340
CCTATACTAT	ATGCTGGTGC	AGATGACCGA	TGAATTTTAT	CCTCAGCTTT	CCGCACATGG	5400
AAAACAGGCG	GTTATTCACG	CCCCCGAGGA	TCTGACCGTG	TCCGGCGACC	CTGATAAACT	5460
CGCGAGAGTC	TTTAACAACA	TTTTGAAAAA	CGCCGCTGCA	TACAGTGAGG	ATAACAGCAT	5520
CATTGACATT	ACCGCGGGCC	TCTCCGGGGA	TGTGGTGTCA	ATCGAATTCA	AGAACACTGG	5580
AAGCATCCCA	AAAGATAAGC	TAGCTGCCAT	ATTTGAAAAG	TTCTATAGGC	TGGACAATGC	5640
TCGTTCTTCC	GATACGGGTG	GCGCGGGACT	TGGATTGGCG	ATTGCAAAAG	AAATTATTGT	5700
TCAGCATGGA	GGGCAGATTT	ACGCGGAAAG	CAATGATAAC	TATACGACGT	TTAGGGTAGA	5760
GCTTCCAGCG	ATGCCAGACT	TGGTTGATAA	AAGGAGGTCC	TAAGAGATGT	ATATAATTTT	5820
TTAGGAAAAT	CTCAAGGTTA	TCTTTACTTT	TTCTTAGGAA	ATTAACAATT	TAATATTAAG	5880
AAACGGCTCG	TTCTTACACG	GTAGACTTAA	TACCGTAAGA	ACGAGCCGTT	TTCGTTCTTC	5940
AGAGAAAGAT	TTGACAAGAT	TACCATTGGC	ATCCCCGTTT	TATTTGGTGC	CTTTCACAGA	6000
AAGGGTTGGT	CTTAATTATG	AATAACATCG	GCATTACTGT	TTATGGATGT	GAGCAGGATG	6060
AGGCAGATGC	ATTCCATGCT	CTTTCGCCTC	GCTTTGGCGT	TATGGCAACG	ATAATTAACG	6120
CCAACGTGTC	GGAATCCAAC	GCCAAATCCG	CGCCTTTCAA	TCAATGTATC	AGTGTGGGAC	6180
ATAAATCAGA	GATTTCCGCC	TCTATTCTTC	TTGCGCTGAA	GAGAGCCGGT	GTGAAATATA	6240
TTTCTACCCG	AAGCATCGGC	TGCAATCATA	TAGATACAAC	TGCTGCTAAG	AGAATGGGCA	6300
TCACTGTCTG	CAATGTGGCG	TACTCGCCGG	ATAGCGTTGC	CGATTATACT	ATGATGCTAA	6360
TTCTTATGGC	AGTACGCAAC	GTAAAATCGA	TTGTGCGCTC	TGTGGAAAAA	CATGATTTCA	6420
GGTTGGACAG	CGACCGTGGC	AAGGTACTCA	GCGACATGAC	AGTTGGTGTG	GTGGGAACGG	6480
GCCAGATAGG	CAAAGCGGTT	ATTGAGCGGC	TGCGAGGATT	TGGATGTAAA	GTGTTGGCTT	6540
ATAGTCGCAG	CCGAAGTATA	GAGGTAAACT	ATGTACCGTT	TGATGAGTTG	CTGCAAAATA	6600
GCGATATCGT	TACGCTTCAT	GTGCCGCTCA	ATACGGATAC	GCACTATATT	ATCAGCCACG	6660
AACAAATACA	GAGAATGAAG	CAAGGAGCAT	TTCTTATCAA	TACTGGGCGC	GGTCCACTTG	6720
TAGATACCTA	TGAGTTGGTT	AAAGCATTAG	AAAACGGGAA	ACTGGGCGGT	GCCGCATTGG	6780
ATGTATTGGA	AGGAGAGGAA	GAGTTTTTCT	ACTCTGATTG	CACCCAAAAA	CCAATTGATA	6840

.ATCAATTTTT	ACTTAAACTT	CAAAGAATGC	CTAACGTGAT	AATCACACCG	CATACGGCCT	6900
ATTATACCGA	GCAAGCGTTG	CGTGATACCG	TTGAAAAAAC	CATTAAAAAC	TGTTTGGATT	6960
TTGAAAGGAG	ACAGGAGCAT	GAATAGAATA	AAAGTTGCAA	TACTGTTTGG	GGGTTGCTCA	7020
GAGGAGCATG	ACGTATCGGT	AAAATCTGCA	ATAGAGATAG	CCGCTAACAT	TAATAAAGAA	7080
AAATACGAGC	CGTTATACAT	TGGAATTACG	AAATCTGGTG	TATGGAAAAT	GTGCGAAAAA	7140
CCTTGCGCGG	AATGGGAAAA	CGACAATTGC	TATTCAGCTG	TACTCTCGCC	GGATAAAAAA	7200
ATGCACGGAT	TACTTGTTAA	AAAGAACCAT	GAATATGAAA	TCAACCATGT	TGATGTAGCA	7260
TTTTTCAGCTT	TGCATGGCAA	GTCAGGTGAA	GATGGATCCA	TACAAGGTCT	GTTTGAATTG	7320
TCCGGTATCC	CTTTTGTAGG	CTGCGATATT	CAAAGCTCAG	CAATTTGTAT	GGACAAATCG	7380
TTGACATACA	TCGTTGCGAA	AAATGCTGGG	ATAGCTACTC	CCGCCTTTTG	GGTTATTAAT	7440
AAAGATGATA	GGCCGGTGGC	AGCTACGTTT	ACCTATCCTG	TTTTTGTAA	GCCGGCGCGT	7500
TCAGGCTCAT	CCTTCGGTGT	GAAAAAAGTC	AATAGCGCGG	ACGAATTGGA	CTACGCAATT	7560
GAATCGGCAA	GACAATATGA	CAGCAAAATC	TTAATTGAGC	AGGCTGTTTC	GGGCTGTGAG	7620
GTCGGTTGTG	CGGTATTGGG	AAACAGTGCC	GCGTTAGTTG	TTGGCGAGGT	GGACCAAATC	7680
AGGCTGCAGT	ACGGAATCTT	TCGTATTTCAT	CAGGAAGTCG	AGCCGGAAAA	AGGCTCTGAA	7740
AACGCAGTTA	TAACCGTTCC	CGCAGACCTT	TCAGCAGAGG	AGCGAGGACG	GATACAGGAA	7800
ACGGCAAAAA	AAATATATAA	AGCGCTCGGC	TGTAGAGGTC	TAGCCCGTGT	GGATATGTTT	7860
TTACAAGATA	ACGGCCGCAT	TGTACTGAAC	GAAGTCAATA	CTCTGCCCGG	TTTCACGTCA	7920
TACAGTCGTT	ATCCCCGTAT	GATGGCCGCT	GCAGGTATTG	CACTTCCCGA	ACTGATTGAC	7980
CGCTTGATCG	TATTAGCGTT	AAAGGGGTGA	TAAGCATGGA	AATAGGATTT	ACTTTTTTTAG	8040
ATGAAATAGT	ACACGGTGTT	CGTTGGGACG	CTAAATATGC	CACTTGGGAT	AATTTACCCG	8100
GAAAACCGGT	TGACGGTTAT	GAAGTAAATC	GCATTGTAGG	GACATACGAG	TTGGCTGAAT	8160
CGCTTTTGAA	GGCAAAAGAA	CTGGCTGCTA	CCCAAGGGTA	CGGATTGCTT	CTATGGGACG	8220
GTTACCGTCC	TAAGCGTGCT	GTAAACTGTT	TTATGCAATG	GGCTGCACAG	CCGGAAAATA	8280
ACCTGACAAA	GGAAAGTTAT	TATCCCAATA	TTGACCGAAC	TGAGATGATT	TCAAAAGGAT	8340
ACGTGGCTTC	AAAATCAAGC	CATAGCCGCG	GCAGTGCCAT	TGATCTTACG	CTTTATCGAT	8400
TAGACACGGG	TGAGCTTGTA	CCAATGGGGA	GCCGATTTGA	TTTTATGGAT	GAACGCTCTC	8460
ATCATGCGGC	AAATGGAATA	TCATGCAATG	AAGCGCAAAA	TCGCAGACGT	TTGCGCTCCA	8520
TCATGGAAAA	CAGTGGGTTT	GAAGCATATA	GCCTCGAATG	GTGGCACTAT	GTATTAAGAG	8580

ACGAACCATA	CCCCAATAGC	TATTTTGATT	TCCCCGTTAA	ATAAACTTTT	AACCGTTGCA	8640
CGGACAAACT	ATATAAGCTA	ACTCTTTCGG	CAGGAAACCC	GACGTATGTA	ACTGGTTCCTT	8700
AGGGAATTTA	TATATAGTAG	ATAGTATTGA	AGATGTAAGG	CAGAGCGATA	TTGCGGTCAT	8760
TATCTGCGTG	CGCTGCGGCA	AGATAGCCTG	ATAATAAGAC	TGATCGCATA	GAGGGGTGGT	8820
ATTTACACACC	GCCCATTTGTC	AACAGGCAGT	TCAGCCTCGT	TAAATTCAGC	ATGGGTATCA	8880
CTTATGAAAA	TTCATCTACA	TTGGTGATAA	TAGTAAATCC	AGTAGGGCGA	AATAATTGAC	8940
TGTAATTTAC	GGGGCAAAAC	GGCACAATCT	CAAACGAGAT	TGTGCCGTTT	AAGGGGAAGA	9000
TTCTAGAAAT	ATTTCATACT	TCCAACATA	TAGTTAAGGA	GGAGACTGAA	AATGAAGAAG	9060
TTGTTTTTTT	TATTGTTATT	GTTATTCTTA	ATATACTTAG	GTTATGACTA	CGTTAATGAA	9120
GCACTGTTTT	CTCAGGAAAA	AGTCGAATTT	CAAATTATG	ATCAAAATCC	CAAAGAACAT	9180
TTAGAAAATA	GTGGGACTTC	TGAAAATACC	CAAGAGAAAA	CAATTACAGA	AGAACAGGTT	9240
TATCAAGGAA	ATCTGCTATT	AATCAATAGT	AAATATCCTG	TTCGCCAAGA	AAGTGTGAAG	9300
TCAGATATCG	TGAATTTATC	TAAACATGAC	GAATTAATAA	ATGGATACGG	GTTGCTTGAT	9360
AGTAATATTT	ATATGTCAAA	AGAAATAGCA	CAAAAATTTT	CAGAGATGGT	CAATGATGCT	9420
GTAAAGGGTG	GCGTTAGTCA	TTTTATTATT	AATAGTGGCT	ATCGAGACTT	TGATGAGCAA	9480
AGTGTGCTTT	ACCAAGAAAT	GGGGGCTGAG	TATGCCTTAC	CAGCAGGTTA	TAGTGAGCAT	9540
AATTCAGGTT	TATCACTAGA	TGTAGGATCA	AGCTTGACGA	AAATGGAACG	AGCCCCTGAA	9600
GGAAAGTGGA	TAGAAGAAAA	TGCTTGGA	TACGGGTCA	TTTTACGTTA	TCCAGAGGAC	9660
AAAACAGAGT	TAACAGGAAT	TCAATATGAA	CCATGGCATA	TTCGCTATGT	TGGTTTACCA	9720
CATAGTGCGA	TTATGAAAGA	AAAGAATTTT	GTTCTCGAGG	AATATATGGA	TTACCTAAAA	9780
GAAGAAAAAA	CCATTTCTGT	TAGTGTAAT	GGGGAAAAAT	ATGAGATCTT	TTATTATCCT	9840
GTTACTAAAA	ATACCACCAT	TCATGTGCCG	ACTAATCTTC	GTTATGAGAT	ATCAGGAAAC	9900
AATATAGACG	GTGTAATTGT	GACAGTGTTT	CCCGGATCAA	CACATACTAA	TTCAAGGAGG	9960
TAAGGATGGC	GGAATGAAAC	CAACGAAATT	AATGAACAGC	ATTATTGTAC	TAGCACTTTT	10020
GGGGTAACGT	TAGCTTTTTA	ATTTAAAACC	CACGTTAAC	AGGACATTGC	TATACTAATG	10080
ATACAACTTA	AACAAAAGAA	TTAGAGGAAA	TTATATTGGG	AAAAATATTA	TCTAGAGGAT	10140
TGCTAGCTTT	ATATTTAGTG	ACACTAATCT	GGTTAGTGTT	ATTCAAATTA	CAATACAATA	10200
TTTTATCAGT	ATTTAATTAT	CATCAAAGAA	GTCTTAACTT	GACTCCATTT	ACTGCTACTG	10260
GGAATTTT	CAGAGATGATA	GATAATGTTA	TAATCTTTAT	TCCATTTGGC	TTGCTTTTGA	10320

ATGTCAATTT TAAAGAAATC GGATTTTAC CTAAGTTTGC TTTTGTACTG GTTTTAAGTC	10380
TTACTTTTGA AATAATTCAA TTTATCTTCG CTATTGGAGC GACAGACATA ACAGATGTAA	10440
TTACAAATAC TGTGAGGC TTTCTTGAC TGAAATTATA TGGTTTAAGC AATAAGCATA	10500
TGAATCAAAA AAAATTAGAC AGAGTTATTA TTTTGTAGG TATACTTTTG CTCGTATTAT	10560
TGCTCGTTTA CCGTACCCAT TTAAGAATAA ATTACGTGTA AGATGTCTAA ATCAAGCAAT	10620
CTGATCTTTC ATACACATAA AGATATTGAA TGAATTGGAT TAGATGGAAA ACGGGATGTG	10680
GGGAAACTCG CCCGTAGGTG TGAAGTGAGG GGAAAACCGG TGATAAAGTA AAAAGCTTAC	10740
CTAACACTAT AGTAACAAAG AAAGCCCAAT TATCAATTTT AGTGCTGAGG AATTGGTCTC	10800
TTTAATAAAT TTCCTTAACG TTGTAAATCC GCATTTTCCT GACGGTACCC C	10851

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2667 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAAGATTACC ATTGGCATCC CCGTTTTATT TGGTGCCTTT CACAGAAAGG GTTGGTCTTA	60
ATTATGAATA ACATCGGCAT TACTGTTTAT GGATGTGAGC AGGATGAGGC AGATGCATTC	120
CATGCTCTTT CGCCTCGCTT TGGCGTTATG GCAACGATAA TTAACGCCAA CGTGTCGGAA	180
TCCAACGCCA AATCCGCGCC TTTCAATCAA TGTATCAGTG TGGGACATAA ATCAGAGATT	240
TCCGCCTCTA TTCTTCTTGC GCTGAAGAGA GCCGGTGTGA AATATATTTT TACCCGAAGC	300
ATCGGCTGCA ATCATATAGA TACAACTGCT GCTAAGAGAA TGGGCATCAC TGTCGACAAT	360
GTGGCGTACT CGCCGGATAG CGTTGCCGAT TATACTATGA TGCTAATTCT TATGGCAGTA	420
CGCAACGTAA AATCGATTGT GCGCTCTGTG GAAAAACATG ATTTTCAGGT GGACAGCGAC	480
CGTGGCAAGG TACTCAGCGA CATGACAGTT GGTGTGGTGG GAACGGGCCA GATAGGCAAA	540
GCGGTTATTG AGCGGCTGCG AGGATTTGGA TGTAAGTGT TGGCTTATAG TCGCAGCCGA	600
AGTATAGAGG TAAACTATGT ACCGTTTGAT GAGTTGATGC AAAATAGCGA TATCGTTACG	660
CTTCATGTGC CGCTCAATAC GGATACGCAC TATATTATCA GCCACGAACA AATACAGAGA	720
ATGAAGCAAG GAGCATTTCT TATCAATACT GGGCGCGGTC CACTTGTAGA TACCTATGAG	780

TTGGTTAAAG	CATTAGAAAA	CGGGAAACTG	GGCGGTGCCG	CATTGGATGT	ATTGGAAGGA	840
GAGGAAGAGT	TTTTCTACTC	TGATTGCACC	CAAAAACCAA	TTGATAATCA	ATTTTTACTT	900
AAACTTCAAA	GAATGCCTAA	CGTGATAATC	ACACCGCATA	CGGCCTATTA	TACCGAGCAA	960
GCGTTGCGTG	ATACCGTTGA	AAAAACCATT	AAAAACTGTT	TGGATTTTGA	AAGGAGACAG	1020
GAGCATGAAT	AGAATAAAAG	TTGCAATACT	GTTTGGGGGT	TGCTCAGAGG	AGCATGACGT	1080
ATCGGTAAAA	TCTGCAATAG	AGATAGCCGC	TAACATTAAT	AAAGAAAAAT	ACGAGCCGTT	1140
ATACATTGGA	ATTACGAAAT	CTGGTGTATG	GAAAATGTGC	GAAAAACCTT	GCGCGGAATG	1200
GGAAAACGAC	AATTGCTATT	CAGCTGTACT	CTCGCCGGAT	AAAAAAATGC	ACGGATTACT	1260
TGTTAAAAAG	AACCATGAAT	ATGAAATCAA	CCATGTTGAT	GTAGCATTTT	CAGCTTTGCA	1320
TGGCAAGTCA	GGTGAAGATG	GATCCATACA	AGGTCTGTTT	GAATTGTCCG	GTATCCCTTT	1380
TGTAGGCTGC	GATATTCAAA	GCTCAGCAAT	TTGTATGGAC	AAATCGTTGA	CATACATCGT	1440
TGCGAAAAAT	GCTGGGATAG	CTACTCCCGC	CTTTTGGGTT	ATTAATAAAG	ATGATAGGCC	1500
GGTGGCAGCT	ACGTTTACCT	ATCCTGTTTT	TGTTAAGCCG	GCGCGTTCAG	GCTCATCCTT	1560
CGGTGTGAAA	AAAGTCAATA	GCGCGGACGA	ATTGGACTAC	GCAATTGAAT	CGGCAAGACA	1620
ATATGACAGC	AAAATCTTAA	TTGAGCAGGC	TGTTTCGGGC	TGTGAGGTCG	GTTGTGCGGT	1680
ATTGGGAAAC	AGTGCCGCGT	TAGTTGTTGG	CGAGGTGGAC	CAAATCAGGC	TGCAGTACGG	1740
AATCTTTCGT	ATTCATCAGG	AAGTCGAGCC	GGAAAAAGGC	TCTGAAAACG	CAGTTATAAC	1800
CGTTCCCGCA	GACCTTTCAG	CAGAGGAGCG	AGGACGGATA	CAGGAAACGG	CAAAAAAAAT	1860
ATATAAAGCG	CTCGGCTGTA	GAGGTCTAGC	CCGTGTGGAT	ATGTTTTTAC	AAGATAACGG	1920
CCGCATTGTA	CTGAACGAAG	TCAATACTCT	GCCCGGTTTC	ACGTCATACA	GTCGTTATCC	1980
CCGTATGATG	GCCGCTGCAG	GTATTGCACT	TCCCGAACTG	ATTGACCGCT	TGATCGTATT	2040
AGCGTTAAAG	GGGTGATAAG	CATGGAAATA	GGATTTACTT	TTTLAGATGA	AATAGTACAC	2100
GGTGTTTCGT	GGGACGCTAA	ATATGCCACT	TGGGATAATT	TCACCGGAAA	ACCGGTTGAC	2160
GGTTATGAAG	TAAATCGCAT	TGTAGGGACA	TACGAGTTGG	CTGAATCGCT	TTTGAAGGCA	2220
AAAGAACTGG	CTGCTACCCA	AGGGTACGGA	TTGCTTCTAT	GGGACGGTTA	CCGTCCTAAG	2280
CGTGCTGTAA	ACTGTTTTAT	GCAATGGGCT	GCACAGCCGG	AAAATAACCT	GACAAAGGAA	2340
AGTTATTATC	CCAATATTGA	CCGAACTGAG	ATGATTTCAA	AAGGATACGT	GGCTTCAAAA	2400
TCAAGCCATA	GCCGCGGCAG	TGCCATTGAT	CTTACGCTTT	ATCGATTAGA	CACGGGTGAG	2460
CTTGTACCAA	TGGGGAGCCG	ATTTGATTTT	ATGGATGAAC	GCTCTCATCA	TGCGGCAAAT	2520

GGAATATCAT GCAATGAAGC GCAAAATCGC AGACGTTTGC GCTCCATCAT GGAAAACAGT	2580
GGGTTTGAAG CATATAGCCT CGAATGGTGG CACTATGTAT TAAGAGACGA ACCATACCCC	2640
AATAGCTATT TTGATTTCCC CGTTAAA	2667

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2964 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2964

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATG AAA ATT GCG AGA GGT AGA GAA TTG CTT ACA CCG GAA CAG AGA CAG	48
Met Lys Ile Ala Arg Gly Arg Glu Leu Leu Thr Pro Glu Gln Arg Gln	
1 5 10 15	
GCT TTT ATG CAA ATC CCT GAA GAT GAA TGG ATA CTG GGG ACC TAC TTC	96
Ala Phe Met Gln Ile Pro Glu Asp Glu Trp Ile Leu Gly Thr Tyr Phe	
20 25 30	
ACT TTT TCC AAA CGG GAT TTA GAA ATA GTT AAT AAG CGA AGG AGG GAA	144
Thr Phe Ser Lys Arg Asp Leu Glu Ile Val Asn Lys Arg Arg Arg Glu	
35 40 45	
GAA AAC CGT TTA GGA TTT GCT GTT CAA TTA GCT GTT CTT CGG TAT CCC	192
Glu Asn Arg Leu Gly Phe Ala Val Gln Leu Ala Val Leu Arg Tyr Pro	
50 55 60	
GGT TGG CCA TAC ACT CAT ATC AAA AGC ATC CCA GAT TCG GTC ATA CAA	240
Gly Trp Pro Tyr Thr His Ile Lys Ser Ile Pro Asp Ser Val Ile Gln	
65 70 75 80	
TAT ATA TCG AAA CAG ATT GGT GTT AGT CCA TCC TCG CTT GAT CAT TAT	288
Tyr Ile Ser Lys Gln Ile Gly Val Ser Pro Ser Ser Leu Asp His Tyr	
85 90 95	
CCT CAA AGG GAA AAT ACA CTT TGG GAT CAT TTG AAA GAA ATT CGA AGT	336
Pro Gln Arg Glu Asn Thr Leu Trp Asp His Leu Lys Glu Ile Arg Ser	
100 105 110	
GAA TAC GAC TTT GTA ACT TTT ACC CTG AGT GAA TAT CGA ATG ACA TTT	384
Glu Tyr Asp Phe Val Thr Phe Thr Leu Ser Glu Tyr Arg Met Thr Phe	
115 120 125	

AAG Lys	TAC Tyr	CTT Leu	CAT His	CAA Gln	TTA Leu	GCT Ala	TTG Leu	GAA Glu	AAT Asn	GGT Gly	GAT Asp	GCC Ala	ATT Ile	CAT His	CTA Leu	432
130						135					140					
CTG Leu	CAT His	GAA Glu	TGC Cys	ATA Ile	GAT Asp	TTT Phe	CTA Leu	AGA Arg	AAA Lys	AAC Asn	AAA Lys	ATT Ile	ATA Ile	CTG Leu	CCT Pro	480
145					150					155					160	
GCT Ala	ATC Ile	ACT Thr	ACA Thr	CTT Leu	GAA Glu	AGA Arg	ATG Met	GTG Val	TGG Trp	GAA Glu	GCA Ala	AGG Arg	GCA Ala	ATG Met	GCT Ala	528
				165					170					175		
GAA Glu	AAG Lys	AAG Lys	CTA Leu	TTT Phe	AAT Asn	ACG Thr	GTT Val	AGT Ser	AAA Lys	TCT Ser	CTA Leu	ACA Thr	AAT Asn	GAG Glu	CAA Gln	576
			180					185					190			
AAA Lys	GAA Glu	AAG Lys	CTT Leu	GAA Glu	GGG Gly	ATT Ile	ATT Ile	ACC Thr	TCG Ser	CAG Gln	CAT His	CCA Pro	TCC Ser	GAA Glu	TCC Ser	624
		195					200					205				
AAT Asn	AAA Lys	ACG Thr	ATA Ile	TTG Leu	GGT Gly	TGG Trp	TTA Leu	AAA Lys	GAG Glu	CCA Pro	CCG Pro	GGT Gly	CAT His	CCT Pro	TCA Ser	672
	210					215					220					
CCC Pro	GAA Glu	ACT Thr	TTT Phe	CTA Leu	AAA Lys	ATA Ile	ATA Ile	GAA Glu	CGA Arg	CTC Leu	GAA Glu	TAC Tyr	ATA Ile	CGA Arg	GGA Gly	720
225					230					235					240	
ATG Met	GAT Asp	TTA Leu	GAA Glu	ACA Thr	GTG Val	CAA Gln	ATT Ile	AGT Ser	CAT His	TTG Leu	CAC His	CGT Arg	AAC Asn	CGC Arg	CTG Leu	768
				245					250					255		
TTG Leu	CAG Gln	CTG Leu	TCT Ser	CGC Arg	TTA Leu	GGC Gly	TCA Ser	AGA Arg	TAC Tyr	GAG Glu	CCG Pro	TAT Tyr	GCA Ala	TTC Phe	CGT Arg	816
			260					265					270			
GAC Asp	TTT Phe	CAA Gln	GAA Glu	AAT Asn	AAA Lys	CGT Arg	TAT Tyr	TCG Ser	ATA Ile	TTA Leu	ACC Thr	ATC Ile	TAT Tyr	TTA Leu	TTA Leu	864
		275					280					285				
CAA Gln	CTT Leu	ACT Thr	CAG Gln	GAG Glu	CTA Leu	ACG Thr	GAT Asp	AAA Lys	GCG Ala	TTT Phe	GAA Glu	ATT Ile	CAT His	GAT Asp	AGG Arg	912
		290				295					300					
CAA Gln	ATA Ile	CTT Leu	AGT Ser	TTG Leu	TTA Leu	TCA Ser	AAA Lys	GGT Gly	CGT Arg	AAG Lys	GCT Ala	CAA Gln	GAG Glu	GAA Glu	ATC Ile	960
					310					315					320	
CAG Gln	AAA Lys	CAA Gln	AAC Asn	GGT Gly	AAA Lys	AAG Lys	CTA Leu	AAT Asn	GAG Glu	AAA Lys	GTT Val	ATA Ile	CAC His	TTT Phe	ACG Thr	1008
				325					330					335		
AAC Asn	ATC Ile	GGA Gly	CAA Gln	GCA Ala	TTA Leu	ATT Ile	AAA Lys	GCA Ala	AGA Arg	GAG Glu	GAA Glu	AAA Lys	TTA Leu	GAC Asp	GTT Val	1056
			340					345					350			

TTT Phe	AAG Lys	GTT Val 355	TTA Leu	GAA Glu	TCG Ser	GTT Val 360	ATT Ile	GAA Glu	TGG Trp	AAT Asn	ACC Thr	TTT Phe 365	GTC Val	TCT Ser	TCA Ser	1104
GTA Val 370	GAA Glu	GAG Glu	GCT Ala	CAG Gln	GAA Glu	CTT Leu 375	GCA Ala	CGT Arg	CCT Pro	GCC Ala	GAC Asp 380	TAT Tyr	GAT Asp	TAT Tyr	TTA Leu	1152
GAC Asp 385	TTA Leu	CTG Leu	CAA Gln	AAA Lys	CGG Arg 390	TTT Phe	TAT Tyr	TCA Ser	CTA Leu	AGA Arg 395	AAA Lys	TAT Tyr	ACG Thr	CCA Pro	ACG Thr 400	1200
CTA Leu	TTA Leu	AGA Arg	GTA Val 405	TTG Leu	GAA Glu	TTT Phe	CAT His	TCT Ser	ACA Thr 410	AAG Lys	GCA Ala	AAT Asn	GAG Glu	CCA Pro 415	CTT Leu	1248
TTA Leu	CAA Gln	GCT Ala 420	GTT Val	GAG Glu	ATT Ile	ATC Ile	CGA Arg	GGA Gly 425	ATG Met	AAC Asn	GAA Glu	TCT Ser	GGA Gly 430	AAG Lys	CGA Arg	1296
AAA Lys	GTG Val 435	CCT Pro	GAT Asp	GAC Asp	TCA Ser	CCT Pro	GTG Val 440	GAT Asp	TTT Phe	ATT Ile	TCA Ser	AAA Lys 445	CGA Arg	TGG Trp	AAA Lys	1344
AGA Arg 450	CAT His	TTA Leu	TAC Tyr	GAG Glu	GAT Asp	GAT Asp 455	GGT Gly	ACA Thr	ACA Thr	ATT Ile	AAT Asn	CGT Arg 460	CAT His	TAC Tyr	TAT Tyr	1392
GAA Glu 465	ATG Met	GCT Ala	GTT Val	TTA Leu	ACA Thr 470	GAA Glu	CTT Leu	CGG Arg	GAG Glu	CAT His 475	GTT Val	CGG Arg	GCA Ala	GGA Gly	GAT Asp 480	1440
GTT Val	TCC Ser	ATT Ile	GTT Val	GGC Gly 485	AGC Ser	AGA Arg	CAA Gln	TAT Tyr	AGG Arg 490	GAT Asp	TTT Phe	GAG Glu	GAA Glu	TAT Tyr 495	TTG Leu	1488
TTT Phe	TCG Ser	GAA Glu 500	GAT Asp	ACA Thr	TGG Trp	AAT Asn	CAA Gln 505	TCG Ser	AAG Lys	GGG Gly	AAT Asn	ACG Thr 510	AGA Arg	TTA Leu	TCA Ser	1536
GTT Val	AGT Ser 515	TTA Leu	TCA Ser	TTC Phe	GAA Glu	GAT Asp	TAT Tyr 520	ATA Ile	ACG Thr	GAG Glu	AGA Arg	ACC Thr 525	AGC Ser	AGC Ser	TTT Phe	1584
AAT Asn 530	GAA Glu	AGG Arg	TTA Leu	AAG Lys	TGG Trp	TTA Leu 535	GCT Ala	GCC Ala	AAT Asn	TCC Ser	AAT Asn 540	AAG Lys	TTA Leu	GAT Asp	GGG Gly	1632
GTT Val 545	TCT Ser	CTT Leu	GAA Glu	AAA Lys	GGA Gly 550	AAG Lys	CTA Leu	TCA Ser	CTT Leu	GCA Ala 555	CGC Arg	TTA Leu	GAA Glu	AAA Lys	GAT Asp 560	1680
GTT Val	CCA Pro	GAA Glu	GAA Glu	GCA Ala 565	AAA Lys	AAA Lys	TTT Phe	AGT Ser	GCA Ala 570	AGC Ser	CTT Leu	TAT Tyr	CAG Gln	ATG Met 575	CTA Leu	1728

CCA	AGA	ATA	AAA	TTA	ACT	GAT	TTA	CTC	ATG	GAT	GTG	GCC	CAT	ATA	ACA	1776
Pro	Arg	Ile	Lys	Leu	Thr	Asp	Leu	Leu	Met	Asp	Val	Ala	His	Ile	Thr	
			580					585					590			
GGA	TTT	CAT	GAG	CAA	TTC	ACT	CAT	GCT	TCC	AAT	AAT	CGA	AAA	CCA	GAT	1824
Gly	Phe	His	Glu	Gln	Phe	Thr	His	Ala	Ser	Asn	Asn	Arg	Lys	Pro	Asp	
		595					600					605				
AAG	GAA	GAA	ACA	ATC	ATT	ATC	ATG	GCT	GCC	CTT	TTA	GGA	ATG	GGA	ATG	1872
Lys	Glu	Glu	Thr	Ile	Ile	Ile	Met	Ala	Ala	Leu	Leu	Gly	Met	Gly	Met	
	610					615					620					
AAT	ATT	GGC	TTG	AGC	AAG	ATG	GCC	GAA	GCC	ACA	CCC	GGA	CTT	ACA	TAT	1920
Asn	Ile	Gly	Leu	Ser	Lys	Met	Ala	Glu	Ala	Thr	Pro	Gly	Leu	Thr	Tyr	
625					630					635					640	
AAG	CAA	CTA	GCC	AAT	GTA	TCT	CAA	TGG	CGC	ATG	TAT	GAA	GAT	GCC	ATG	1968
Lys	Gln	Leu	Ala	Asn	Val	Ser	Gln	Trp	Arg	Met	Tyr	Glu	Asp	Ala	Met	
				645					650					655		
AAT	AAA	GCC	CAA	GCC	ATA	TTA	GTA	AAC	TTT	CAT	CAT	AAA	TTA	CAA	TTG	2016
Asn	Lys	Ala	Gln	Ala	Ile	Leu	Val	Asn	Phe	His	His	Lys	Leu	Gln	Leu	
			660					665					670			
CCT	TTC	TAT	TGG	GGC	GAC	GGT	ACA	ACA	TCT	TCG	TCA	GAT	GGT	ATG	AGA	2064
Pro	Phe	Tyr	Trp	Gly	Asp	Gly	Thr	Thr	Ser	Ser	Ser	Asp	Gly	Met	Arg	
		675					680					685				
ATG	CAG	CTA	GGT	GTT	TCA	TCA	CTA	CAT	GCA	GAT	GCA	AAT	CCA	CAT	TAT	2112
Met	Gln	Leu	Gly	Val	Ser	Ser	Leu	His	Ala	Asp	Ala	Asn	Pro	His	Tyr	
	690					695					700					
GGA	ACT	GGA	AAA	GGA	GCC	ACC	ATC	TAC	CGA	TTT	ACA	AGT	GAT	CAA	TTC	2160
Gly	Thr	Gly	Lys	Gly	Ala	Thr	Ile	Tyr	Arg	Phe	Thr	Ser	Asp	Gln	Phe	
705					710					715					720	
TCT	TCT	TAC	TAC	ACA	AAG	ATT	ATT	CAT	ACT	AAT	TCA	AGA	GAT	GCG	ATT	2208
Ser	Ser	Tyr	Tyr	Thr	Lys	Ile	Ile	His	Thr	Asn	Ser	Arg	Asp	Ala	Ile	
				725					730					735		
CAT	GTT	TTG	GAT	GGT	TTG	TTA	CAT	CAT	GAG	ACG	GAT	CTA	AAC	ATA	GAG	2256
His	Val	Leu	Asp	Gly	Leu	Leu	His	His	Glu	Thr	Asp	Leu	Asn	Ile	Glu	
			740					745					750			
GAA	CAT	TAT	ACA	GAC	ACT	GCC	GGT	TAC	ACT	GAC	CAA	ATA	TTC	GGA	CTG	2304
Glu	His	Tyr	Thr	Asp	Thr	Ala	Gly	Tyr	Thr	Asp	Gln	Ile	Phe	Gly	Leu	
		755					760					765				
ACT	CAT	TTA	TTA	GGA	TTT	AAA	TTT	GCC	CCA	AGA	ATA	AGG	GAT	TTA	TCG	2352
Thr	His	Leu	Leu	Gly	Phe	Lys	Phe	Ala	Pro	Arg	Ile	Arg	Asp	Leu	Ser	
	770					775					780					
GAC	TCA	AAA	TTA	TTT	ACG	ATA	GAT	AAA	GCA	AGT	GAG	TAT	CCA	AAA	CTA	2400
Asp	Ser	Lys	Leu	Phe	Thr	Ile	Asp	Lys	Ala	Ser	Glu	Tyr	Pro	Lys	Leu	
785					790					795					800	

GAA	GCC	ATT	TTA	CGT	GGA	CAA	ATA	AAT	ACA	AAG	GTC	ATT	AAA	GAA	AAT	2448
Glu	Ala	Ile	Leu	Arg	Gly	Gln	Ile	Asn	Thr	Lys	Val	Ile	Lys	Glu	Asn	
				805					810					815		
TAT	GAG	GAT	GTT	TTG	CGA	TTA	GCT	CAT	TCT	ATA	AGG	GAG	GGA	ACA	GTT	2496
Tyr	Glu	Asp	Val	Leu	Arg	Leu	Ala	His	Ser	Ile	Arg	Glu	Gly	Thr	Val	
			820					825					830			
TCA	GCA	TCC	CTT	ATT	ATG	GGG	AAG	CTA	GGT	TCC	TAT	TCA	AGA	CAA	AAC	2544
Ser	Ala	Ser	Leu	Ile	Met	Gly	Lys	Leu	Gly	Ser	Tyr	Ser	Arg	Gln	Asn	
		835					840					845				
AGC	TTA	GCT	ACA	GCC	TTA	CGT	GAG	ATG	GGC	CGA	ATA	GAA	AAA	ACG	ATC	2592
Ser	Leu	Ala	Thr	Ala	Leu	Arg	Glu	Met	Gly	Arg	Ile	Glu	Lys	Thr	Ile	
	850					855					860					
TTT	ATT	TTG	AAT	TAT	ATA	TCG	GAT	GAA	TCA	TTA	AGA	AGA	AAA	ATA	CAA	2640
Phe	Ile	Leu	Asn	Tyr	Ile	Ser	Asp	Glu	Ser	Leu	Arg	Arg	Lys	Ile	Gln	
865					870					875					880	
AGA	GGA	TTG	AAT	AAA	GGA	GAA	GCC	ATG	AAT	GGA	TTG	GCA	AGA	GCT	ATT	2688
Arg	Gly	Leu	Asn	Lys	Gly	Glu	Ala	Met	Asn	Gly	Leu	Ala	Arg	Ala	Ile	
				885					890					895		
TTC	TTC	GGA	AAA	CAA	GGT	GAG	CTT	AGA	GAA	CGC	ACC	ATA	CAG	CAT	CAA	2736
Phe	Phe	Gly	Lys	Gln	Gly	Glu	Leu	Arg	Glu	Arg	Thr	Ile	Gln	His	Gln	
			900					905					910			
TTG	CAA	AGA	GCC	AGT	GCT	TTA	AAC	ATA	ATT	ATC	AAT	GCT	ATA	AGT	ATT	2784
Leu	Gln	Arg	Ala	Ser	Ala	Leu	Asn	Ile	Ile	Ile	Asn	Ala	Ile	Ser	Ile	
		915					920					925				
TGG	AAT	ACT	CTC	CAC	CTA	ACA	ACA	GCA	GTT	GAA	TAT	AAA	AAA	CGG	ACA	2832
Trp	Asn	Thr	Leu	His	Leu	Thr	Thr	Ala	Val	Glu	Tyr	Lys	Lys	Arg	Thr	
	930					935					940					
GGT	AGC	TTT	AAT	GAA	GAT	TTG	TTA	CAC	CAT	ATG	TCG	CCC	TTA	GGT	TGG	2880
Gly	Ser	Phe	Asn	Glu	Asp	Leu	Leu	His	His	Met	Ser	Pro	Leu	Gly	Trp	
945					950					955				960		
GAA	CAT	ATT	AAT	TTA	CTA	GGA	GAA	TAC	CAT	TTT	AAC	TCA	GAG	AAA	GTA	2928
Glu	His	Ile	Asn	Leu	Leu	Gly	Glu	Tyr	His	Phe	Asn	Ser	Glu	Lys	Val	
			965					970						975		
GTC	TCA	TTA	AAT	TCT	TTA	AGA	CCA	CTA	AAA	CTT	TCT					2964
Val	Ser	Leu	Asn	Ser	Leu	Arg	Pro	Leu	Lys	Leu	Ser					
			980					985								

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 988 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met	Lys	Ile	Ala	Arg	Gly	Arg	Glu	Leu	Leu	Thr	Pro	Glu	Gln	Arg	Gln	
1				5				10						15		
Ala	Phe	Met	Gln	Ile	Pro	Glu	Asp	Glu	Trp	Ile	Leu	Gly	Thr	Tyr	Phe	
			20					25					30			
Thr	Phe	Ser	Lys	Arg	Asp	Leu	Glu	Ile	Val	Asn	Lys	Arg	Arg	Arg	Glu	
		35					40					45				
Glu	Asn	Arg	Leu	Gly	Phe	Ala	Val	Gln	Leu	Ala	Val	Leu	Arg	Tyr	Pro	
	50					55					60					
Gly	Trp	Pro	Tyr	Thr	His	Ile	Lys	Ser	Ile	Pro	Asp	Ser	Val	Ile	Gln	
65					70					75					80	
Tyr	Ile	Ser	Lys	Gln	Ile	Gly	Val	Ser	Pro	Ser	Ser	Leu	Asp	His	Tyr	
				85					90					95		
Pro	Gln	Arg	Glu	Asn	Thr	Leu	Trp	Asp	His	Leu	Lys	Glu	Ile	Arg	Ser	
			100					105					110			
Glu	Tyr	Asp	Phe	Val	Thr	Phe	Thr	Leu	Ser	Glu	Tyr	Arg	Met	Thr	Phe	
		115					120					125				
Lys	Tyr	Leu	His	Gln	Leu	Ala	Leu	Glu	Asn	Gly	Asp	Ala	Ile	His	Leu	
	130					135					140					
Leu	His	Glu	Cys	Ile	Asp	Phe	Leu	Arg	Lys	Asn	Lys	Ile	Ile	Leu	Pro	
145					150					155					160	
Ala	Ile	Thr	Thr	Leu	Glu	Arg	Met	Val	Trp	Glu	Ala	Arg	Ala	Met	Ala	
				165					170					175		
Glu	Lys	Lys	Leu	Phe	Asn	Thr	Val	Ser	Lys	Ser	Leu	Thr	Asn	Glu	Gln	
			180					185					190			
Lys	Glu	Lys	Leu	Glu	Gly	Ile	Ile	Thr	Ser	Gln	His	Pro	Ser	Glu	Ser	
	195					200						205				
Asn	Lys	Thr	Ile	Leu	Gly	Trp	Leu	Lys	Glu	Pro	Pro	Gly	His	Pro	Ser	
	210					215					220					
Pro	Glu	Thr	Phe	Leu	Lys	Ile	Ile	Glu	Arg	Leu	Glu	Tyr	Ile	Arg	Gly	
225					230					235					240	
Met	Asp	Leu	Glu	Thr	Val	Gln	Ile	Ser	His	Leu	His	Arg	Asn	Arg	Leu	
				245					250					255		
Leu	Gln	Leu	Ser	Arg	Leu	Gly	Ser	Arg	Tyr	Glu	Pro	Tyr	Ala	Phe	Arg	
			260					265					270			
Asp	Phe	Gln	Glu	Asn	Lys	Arg	Tyr	Ser	Ile	Leu	Thr	Ile	Tyr	Leu	Leu	
	275						280					285				
Gln	Leu	Thr	Gln	Glu	Leu	Thr	Asp	Lys	Ala	Phe	Glu	Ile	His	Asp	Arg	
	290					295					300					

Gln	Ile	Leu	Ser	Leu	Leu	Ser	Lys	Gly	Arg	Lys	Ala	Gln	Glu	Glu	Ile	305	310	315	320
Gln	Lys	Gln	Asn	Gly	Lys	Lys	Leu	Asn	Glu	Lys	Val	Ile	His	Phe	Thr	325	330	335	
Asn	Ile	Gly	Gln	Ala	Leu	Ile	Lys	Ala	Arg	Glu	Glu	Lys	Leu	Asp	Val	340	345	350	
Phe	Lys	Val	Leu	Glu	Ser	Val	Ile	Glu	Trp	Asn	Thr	Phe	Val	Ser	Ser	355	360	365	
Val	Glu	Glu	Ala	Gln	Glu	Leu	Ala	Arg	Pro	Ala	Asp	Tyr	Asp	Tyr	Leu	370	375	380	
Asp	Leu	Leu	Gln	Lys	Arg	Phe	Tyr	Ser	Leu	Arg	Lys	Tyr	Thr	Pro	Thr	385	390	395	400
Leu	Leu	Arg	Val	Leu	Glu	Phe	His	Ser	Thr	Lys	Ala	Asn	Glu	Pro	Leu	405	410	415	
Leu	Gln	Ala	Val	Glu	Ile	Ile	Arg	Gly	Met	Asn	Glu	Ser	Gly	Lys	Arg	420	425	430	
Lys	Val	Pro	Asp	Asp	Ser	Pro	Val	Asp	Phe	Ile	Ser	Lys	Arg	Trp	Lys	435	440	445	
Arg	His	Leu	Tyr	Glu	Asp	Asp	Gly	Thr	Thr	Ile	Asn	Arg	His	Tyr	Tyr	450	455	460	
Glu	Met	Ala	Val	Leu	Thr	Glu	Leu	Arg	Glu	His	Val	Arg	Ala	Gly	Asp	465	470	475	480
Val	Ser	Ile	Val	Gly	Ser	Arg	Gln	Tyr	Arg	Asp	Phe	Glu	Glu	Tyr	Leu	485	490	495	
Phe	Ser	Glu	Asp	Thr	Trp	Asn	Gln	Ser	Lys	Gly	Asn	Thr	Arg	Leu	Ser	500	505	510	
Val	Ser	Leu	Ser	Phe	Glu	Asp	Tyr	Ile	Thr	Glu	Arg	Thr	Ser	Ser	Phe	515	520	525	
Asn	Glu	Arg	Leu	Lys	Trp	Leu	Ala	Ala	Asn	Ser	Asn	Lys	Leu	Asp	Gly	530	535	540	
Val	Ser	Leu	Glu	Lys	Gly	Lys	Leu	Ser	Leu	Ala	Arg	Leu	Glu	Lys	Asp	545	550	555	560
Val	Pro	Glu	Glu	Ala	Lys	Lys	Phe	Ser	Ala	Ser	Leu	Tyr	Gln	Met	Leu	565	570	575	
Pro	Arg	Ile	Lys	Leu	Thr	Asp	Leu	Leu	Met	Asp	Val	Ala	His	Ile	Thr	580	585	590	
Gly	Phe	His	Glu	Gln	Phe	Thr	His	Ala	Ser	Asn	Asn	Arg	Lys	Pro	Asp	595	600	605	

Lys	Glu	Glu	Thr	Ile	Ile	Ile	Met	Ala	Ala	Leu	Leu	Gly	Met	Gly	Met		
610						615					620						
Asn	Ile	Gly	Leu	Ser	Lys	Met	Ala	Glu	Ala	Thr	Pro	Gly	Leu	Thr	Tyr		
625					630					635					640		
Lys	Gln	Leu	Ala	Asn	Val	Ser	Gln	Trp	Arg	Met	Tyr	Glu	Asp	Ala	Met		
				645					650					655			
Asn	Lys	Ala	Gln	Ala	Ile	Leu	Val	Asn	Phe	His	His	Lys	Leu	Gln	Leu		
			660					665					670				
Pro	Phe	Tyr	Trp	Gly	Asp	Gly	Thr	Thr	Ser	Ser	Ser	Asp	Gly	Met	Arg		
		675					680					685					
Met	Gln	Leu	Gly	Val	Ser	Ser	Leu	His	Ala	Asp	Ala	Asn	Pro	His	Tyr		
		690				695					700						
Gly	Thr	Gly	Lys	Gly	Ala	Thr	Ile	Tyr	Arg	Phe	Thr	Ser	Asp	Gln	Phe		
705					710					715					720		
Ser	Ser	Tyr	Tyr	Thr	Lys	Ile	Ile	His	Thr	Asn	Ser	Arg	Asp	Ala	Ile		
				725					730					735			
His	Val	Leu	Asp	Gly	Leu	Leu	His	His	Glu	Thr	Asp	Leu	Asn	Ile	Glu		
			740					745					750				
Glu	His	Tyr	Thr	Asp	Thr	Ala	Gly	Tyr	Thr	Asp	Gln	Ile	Phe	Gly	Leu		
		755					760					765					
Thr	His	Leu	Leu	Gly	Phe	Lys	Phe	Ala	Pro	Arg	Ile	Arg	Asp	Leu	Ser		
						775					780						
Asp	Ser	Lys	Leu	Phe	Thr	Ile	Asp	Lys	Ala	Ser	Glu	Tyr	Pro	Lys	Leu		
785					790					795					800		
Glu	Ala	Ile	Leu	Arg	Gly	Gln	Ile	Asn	Thr	Lys	Val	Ile	Lys	Glu	Asn		
				805					810					815			
Tyr	Glu	Asp	Val	Leu	Arg	Leu	Ala	His	Ser	Ile	Arg	Glu	Gly	Thr	Val		
			820					825					830				
Ser	Ala	Ser	Leu	Ile	Met	Gly	Lys	Leu	Gly	Ser	Tyr	Ser	Arg	Gln	Asn		
		835					840					845					
Ser	Leu	Ala	Thr	Ala	Leu	Arg	Glu	Met	Gly	Arg	Ile	Glu	Lys	Thr	Ile		
		850				855					860						
Phe	Ile	Leu	Asn	Tyr	Ile	Ser	Asp	Glu	Ser	Leu	Arg	Arg	Lys	Ile	Gln		
865					870					875					880		
Arg	Gly	Leu	Asn	Lys	Gly	Glu	Ala	Met	Asn	Gly	Leu	Ala	Arg	Ala	Ile		
			885						890					895			
Phe	Phe	Gly	Lys	Gln	Gly	Glu	Leu	Arg	Glu	Arg	Thr	Ile	Gln	His	Gln		
			900					905					910				

GAT TTA TCA GAA GAT AAT CCA TAC AGC CAA TTC TTA ATT ACT GTA ATG	336
Asp Leu Ser Glu Asp Asn Pro Tyr Ser Gln Phe Leu Ile Thr Val Met	
100 105 110	
GCT GGT GTT AAC CAA TTA GAG CGA GAT CTT ATT CGG ATG AGA CAA CGT	384
Ala Gly Val Asn Gln Leu Glu Arg Asp Leu Ile Arg Met Arg Gln Arg	
115 120 125	
GAA GGG ATT GAA TTG GCT AAG AAA GAA GGA AAG TTT AAA GGT CGA TTA	432
Glu Gly Ile Glu Leu Ala Lys Lys Glu Gly Lys Phe Lys Gly Arg Leu	
130 135 140	
AAG AAG TAT CAT AAA AAT CAC GCA GGA ATG AAT TAT GCG GTA AAG CTA	480
Lys Lys Tyr His Lys Asn His Ala Gly Met Asn Tyr Ala Val Lys Leu	
145 150 155 160	
TAT AAA GAA GGA AAT ATG ACT GTA AAT CAA ATT TGT GAA ATT ACT AAT	528
Tyr Lys Glu Gly Asn Met Thr Val Asn Gln Ile Cys Glu Ile Thr Asn	
165 170 175	
GTA TCT AGG GCT TCA TTA TAC AGG AAA TTA TCA GAA GTG AAT AAT	573
Val Ser Arg Ala Ser Leu Tyr Arg Lys Leu Ser Glu Val Asn Asn	
180 185 190	

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 191 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Leu Arg Lys Ile Gly Tyr Ile Arg Val Ser Ser Thr Asn Gln Asn Pro	
1 5 10 15	
Ser Arg Gln Phe Gln Gln Leu Asn Glu Ile Gly Met Asp Ile Ile Tyr	
20 25 30	
Glu Glu Lys Val Ser Gly Ala Thr Lys Asp Arg Glu Gln Leu Gln Lys	
35 40 45	
Val Leu Asp Asp Leu Gln Glu Asp Asp Ile Ile Tyr Val Thr Asp Leu	
50 55 60	
Thr Arg Ile Thr Arg Ser Thr Gln Asp Leu Phe Glu Leu Ile Asp Asn	
65 70 75 80	
Ile Arg Asp Lys Lys Ala Ser Leu Lys Ser Leu Lys Asp Thr Trp Leu	
85 90 95	
Asp Leu Ser Glu Asp Asn Pro Tyr Ser Gln Phe Leu Ile Thr Val Met	
100 105 110	

Ala Gly Val Asn Gln Leu Glu Arg Asp Leu Ile Arg Met Arg Gln Arg
115 120 125

Glu Gly Ile Glu Leu Ala Lys Lys Glu Gly Lys Phe Lys Gly Arg Leu
130 135 140

Lys Lys Tyr His Lys Asn His Ala Gly Met Asn Tyr Ala Val Lys Leu
145 150 155 160

Tyr Lys Glu Gly Asn Met Thr Val Asn Gln Ile Cys Glu Ile Thr Asn
165 170 175

Val Ser Arg Ala Ser Leu Tyr Arg Lys Leu Ser Glu Val Asn Asn
180 185 190

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 909 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..909

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATG AAG AAG TTG TTT TTT TTA TTG TTA TTG TTA TTC TTA ATA TAC TTA	48
Met Lys Lys Leu Phe Phe Leu Leu Leu Leu Phe Leu Ile Tyr Leu	
1 5 10 15	
GGT TAT GAC TAC GTT AAT GAA GCA CTG TTT TCT CAG GAA AAA GTC GAA	96
Gly Tyr Asp Tyr Val Asn Glu Ala Leu Phe Ser Gln Glu Lys Val Glu	
20 25 30	
TTT CAA AAT TAT GAT CAA AAT CCC AAA GAA CAT TTA GAA AAT AGT GGG	144
Phe Gln Asn Tyr Asp Gln Asn Pro Lys Glu His Leu Glu Asn Ser Gly	
35 40 45	
ACT TCT GAA AAT ACC CAA GAG AAA ACA ATT ACA GAA GAA CAG GTT TAT	192
Thr Ser Glu Asn Thr Gln Glu Lys Thr Ile Thr Glu Glu Gln Val Tyr	
50 55 60	
CAA GGA AAT CTG CTA TTA ATC AAT AGT AAA TAT CCT GTT CGC CAA GAA	240
Gln Gly Asn Leu Leu Ile Asn Ser Lys Tyr Pro Val Arg Gln Glu	
65 70 75 80	
AGT GTG AAG TCA GAT ATC GTG AAT TTA TCT AAA CAT GAC GAA TTA ATA	288
Ser Val Lys Ser Asp Ile Val Asn Leu Ser Lys His Asp Glu Leu Ile	
85 90 95	

AAT GGA TAC GGG TTG CTT GAT AGT AAT ATT TAT ATG TCA AAA GAA ATA	336
Asn Gly Tyr Gly Leu Leu Asp Ser Asn Ile Tyr Met Ser Lys Glu Ile	
100 105 110	
GCA CAA AAA TTT TCA GAG ATG GTC AAT GAT GCT GTA AAG GGT GGC GTT	384
Ala Gln Lys Phe Ser Glu Met Val Asn Asp Ala Val Lys Gly Gly Val	
115 120 125	
AGT CAT TTT ATT ATT AAT AGT GGC TAT CGA GAC TTT GAT GAG CAA AGT	432
Ser His Phe Ile Ile Asn Ser Gly Tyr Arg Asp Phe Asp Glu Gln Ser	
130 135 140	
GTG CTT TAC CAA GAA ATG GGG GCT GAG TAT GCC TTA CCA GCA GGT TAT	480
Val Leu Tyr Gln Glu Met Gly Ala Glu Tyr Ala Leu Pro Ala Gly Tyr	
145 150 155 160	
AGT GAG CAT AAT TCA GGT TTA TCA CTA GAT GTA GGA TCA AGC TTG ACG	528
Ser Glu His Asn Ser Gly Leu Ser Leu Asp Val Gly Ser Ser Leu Thr	
165 170 175	
AAA ATG GAA CGA GCC CCT GAA GGA AAG TGG ATA GAA GAA AAT GCT TGG	576
Lys Met Glu Arg Ala Pro Glu Gly Lys Trp Ile Glu Glu Asn Ala Trp	
180 185 190	
AAA TAC GGG TTC ATT TTA CGT TAT CCA GAG GAC AAA ACA GAG TTA ACA	624
Lys Tyr Gly Phe Ile Leu Arg Tyr Pro Glu Asp Lys Thr Glu Leu Thr	
195 200 205	
GGA ATT CAA TAT GAA CCA TGG CAT ATT CGC TAT GTT GGT TTA CCA CAT	672
Gly Ile Gln Tyr Glu Pro Trp His Ile Arg Tyr Val Gly Leu Pro His	
210 215 220	
AGT GCG ATT ATG AAA GAA AAG AAT TTC GTT CTC GAG GAA TAT ATG GAT	720
Ser Ala Ile Met Lys Glu Lys Asn Phe Val Leu Glu Glu Tyr Met Asp	
225 230 235 240	
TAC CTA AAA GAA GAA AAA ACC ATT TCT GTT AGT GTA AAT GGG GAA AAA	768
Tyr Leu Lys Glu Glu Lys Thr Ile Ser Val Ser Val Asn Gly Glu Lys	
245 250 255	
TAT GAG ATC TTT TAT TAT CCT GTT ACT AAA AAT ACC ACC ATT CAT GTG	816
Tyr Glu Ile Phe Tyr Tyr Pro Val Thr Lys Asn Thr Thr Ile His Val	
260 265 270	
CCG ACT AAT CTT CGT TAT GAG ATA TCA GGA AAC AAT ATA GAC GGT GTA	864
Pro Thr Asn Leu Arg Tyr Glu Ile Ser Gly Asn Asn Ile Asp Gly Val	
275 280 285	
ATT GTG ACA GTG TTT CCC GGA TCA ACA CAT ACT AAT TCA AGG AGG	909
Ile Val Thr Val Phe Pro Gly Ser Thr His Thr Asn Ser Arg Arg	
290 295 300	

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 303 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met	Lys	Lys	Leu	Phe	Phe	Leu	Leu	Leu	Leu	Leu	Phe	Leu	Ile	Tyr	Leu	
1				5				10						15		
Gly	Tyr	Asp	Tyr	Val	Asn	Glu	Ala	Leu	Phe	Ser	Gln	Glu	Lys	Val	Glu	
		20						25					30			
Phe	Gln	Asn	Tyr	Asp	Gln	Asn	Pro	Lys	Glu	His	Leu	Glu	Asn	Ser	Gly	
		35					40					45				
Thr	Ser	Glu	Asn	Thr	Gln	Glu	Lys	Thr	Ile	Thr	Glu	Glu	Gln	Val	Tyr	
	50					55					60					
Gln	Gly	Asn	Leu	Leu	Leu	Ile	Asn	Ser	Lys	Tyr	Pro	Val	Arg	Gln	Glu	
65					70					75					80	
Ser	Val	Lys	Ser	Asp	Ile	Val	Asn	Leu	Ser	Lys	His	Asp	Glu	Leu	Ile	
				85					90					95		
Asn	Gly	Tyr	Gly	Leu	Leu	Asp	Ser	Asn	Ile	Tyr	Met	Ser	Lys	Glu	Ile	
			100					105					110			
Ala	Gln	Lys	Phe	Ser	Glu	Met	Val	Asn	Asp	Ala	Val	Lys	Gly	Gly	Val	
		115					120					125				
Ser	His	Phe	Ile	Ile	Asn	Ser	Gly	Tyr	Arg	Asp	Phe	Asp	Glu	Gln	Ser	
	130					135					140					
Val	Leu	Tyr	Gln	Glu	Met	Gly	Ala	Glu	Tyr	Ala	Leu	Pro	Ala	Gly	Tyr	
145					150					155					160	
Ser	Glu	His	Asn	Ser	Gly	Leu	Ser	Leu	Asp	Val	Gly	Ser	Ser	Leu	Thr	
				165					170					175		
Lys	Met	Glu	Arg	Ala	Pro	Glu	Gly	Lys	Trp	Ile	Glu	Glu	Asn	Ala	Trp	
			180					185					190			
Lys	Tyr	Gly	Phe	Ile	Leu	Arg	Tyr	Pro	Glu	Asp	Lys	Thr	Glu	Leu	Thr	
		195					200					205				
Gly	Ile	Gln	Tyr	Glu	Pro	Trp	His	Ile	Arg	Tyr	Val	Gly	Leu	Pro	His	
	210					215					220					
Ser	Ala	Ile	Met	Lys	Glu	Lys	Asn	Phe	Val	Leu	Glu	Glu	Tyr	Met	Asp	
225					230					235					240	
Tyr	Leu	Lys	Glu	Glu	Lys	Thr	Ile	Ser	Val	Ser	Val	Asn	Gly	Glu	Lys	
				245					250					255		

Tyr Glu Ile Phe Tyr Tyr Pro Val Thr Lys Asn Thr Thr Ile His Val
260 265 270

Pro Thr Asn Leu Arg Tyr Glu Ile Ser Gly Asn Asn Ile Asp Gly Val
275 280 285

Ile Val Thr Val Phe Pro Gly Ser Thr His Thr Asn Ser Arg Arg
290 295 300

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..483

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTG GGA AAA ATA TTA TCT AGA GGA TTG CTA GCT TTA TAT TTA GTG ACA	48
Leu Gly Lys Ile Leu Ser Arg Gly Leu Leu Ala Leu Tyr Leu Val Thr	
1 5 10 15	
CTA ATC TGG TTA GTG TTA TTC AAA TTA CAA TAC AAT ATT TTA TCA GTA	96
Leu Ile Trp Leu Val Leu Phe Lys Leu Gln Tyr Asn Ile Leu Ser Val	
20 25 30	
TTT AAT TAT CAT CAA AGA AGT CTT AAC TTG ACT CCA TTT ACT GCT ACT	144
Phe Asn Tyr His Gln Arg Ser Leu Asn Leu Thr Pro Phe Thr Ala Thr	
35 40 45	
GGG AAT TTC AGA GAG ATG ATA GAT AAT GTT ATA ATC TTT ATT CCA TTT	192
Gly Asn Phe Arg Glu Met Ile Asp Asn Val Ile Ile Phe Ile Pro Phe	
50 55 60	
GGC TTG CTT TTG AAT GTC AAT TTT AAA GAA ATC GGA TTT TTA CCT AAG	240
Gly Leu Leu Leu Asn Val Asn Phe Lys Glu Ile Gly Phe Leu Pro Lys	
65 70 75 80	
TTT GCT TTT GTA CTG GTT TTA AGT CTT ACT TTT GAA ATA ATT CAA TTT	288
Phe Ala Phe Val Leu Val Leu Ser Leu Thr Phe Glu Ile Ile Gln Phe	
85 90 95	
ATC TTC GCT ATT GGA GCG ACA GAC ATA ACA GAT GTA ATT ACA AAT ACT	336
Ile Phe Ala Ile Gly Ala Thr Asp Ile Thr Asp Val Ile Thr Asn Thr	
100 105 110	
GTT GGA GGC TTT CTT GGA CTG AAA TTA TAT GGT TTA AGC AAT AAG CAT	384
Val Gly Gly Phe Leu Gly Leu Lys Leu Tyr Gly Leu Ser Asn Lys His	
115 120 125	

ATG AAT CAA AAA AAA TTA GAC AGA GTT ATT ATT TTT GTA GGT ATA CTT	432
Met Asn Gln Lys Lys Leu Asp Arg Val Ile Ile Phe Val Gly Ile Leu	
130 135 140	
TTG CTC GTA TTA TTG CTC GTT TAC CGT ACC CAT TTA AGA ATA AAT TAC	480
Leu Leu Val Leu Leu Leu Val Tyr Arg Thr His Leu Arg Ile Asn Tyr	
145 150 155 160	
GTG	483
Val	

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Leu Gly Lys Ile Leu Ser Arg Gly Leu Leu Ala Leu Tyr Leu Val Thr	
1 5 10 15	
Leu Ile Trp Leu Val Leu Phe Lys Leu Gln Tyr Asn Ile Leu Ser Val	
20 25 30	
Phe Asn Tyr His Gln Arg Ser Leu Asn Leu Thr Pro Phe Thr Ala Thr	
35 40 45	
Gly Asn Phe Arg Glu Met Ile Asp Asn Val Ile Ile Phe Ile Pro Phe	
50 55 60	
Gly Leu Leu Leu Asn Val Asn Phe Lys Glu Ile Gly Phe Leu Pro Lys	
65 70 75 80	
Phe Ala Phe Val Leu Val Leu Ser Leu Thr Phe Glu Ile Ile Gln Phe	
85 90 95	
Ile Phe Ala Ile Gly Ala Thr Asp Ile Thr Asp Val Ile Thr Asn Thr	
100 105 110	
Val Gly Gly Phe Leu Gly Leu Lys Leu Tyr Gly Leu Ser Asn Lys His	
115 120 125	
Met Asn Gln Lys Lys Leu Asp Arg Val Ile Ile Phe Val Gly Ile Leu	
130 135 140	
Leu Leu Val Leu Leu Leu Val Tyr Arg Thr His Leu Arg Ile Asn Tyr	
145 150 155 160	
Val	

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATGCTCCTGT CTCCTTTC

18

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2296 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Lys Leu Phe Phe Leu Leu Ile Cys Arg Phe Thr Asn Arg Ile Lys Leu
1 5 10 15

Leu Phe Ser His Cys Pro Cys Phe Pro His His Ser Phe Lys Cys Ser
20 25 30

Asp Ser Arg Gln Tyr Asn Phe Val Phe Ser Lys Ile Tyr Ala Phe Met
35 40 45

Gln Met Asn Gly Ile Thr Ile Phe Gln Ser Leu Met Lys Val Leu Lys
50 55 60

Cys His Ser Ile Phe Thr Gln Gly Lys Ser Tyr Lys Val Val Phe Thr
65 70 75 80

Ser Asn Phe Phe Gln Met Ile Pro Lys Cys Ile Phe Pro Leu Arg Ile
85 90 95

Met Ile Lys Arg Gly Trp Thr Asn Thr Asn Leu Phe Arg Tyr Ile Leu
100 105 110

Tyr Asp Arg Ile Trp Asp Ala Phe Asp Met Ser Val Trp Pro Thr Gly
115 120 125

Ile Pro Lys Asn Ser Leu Asn Ser Lys Ser Thr Val Phe Phe Pro Pro
130 135 140

Ser Leu Ile Asn Tyr Phe Ile Pro Phe Gly Lys Ser Glu Val Gly Pro
145 150 155 160

Gln	Tyr	Pro	Phe	Ile	Phe	Arg	Asp	Leu	His	Lys	Ser	Leu	Ser	Leu	Phe	165	170	175
Arg	Cys	Lys	Gln	Phe	Ser	Thr	Ser	Arg	Asn	Phe	His	Ser	Val	Ser	Phe	180	185	190
His	Phe	Cys	Ile	Phe	Asn	Leu	Leu	Val	Gln	Leu	Tyr	Ile	Asn	Arg	Val	195	200	205
Tyr	Ser	Ile	Asp	Thr	Asn	Val	Val	Asp	Asn	His	Ser	Glu	Arg	Leu	Ile	210	215	220
Arg	Leu	Val	Ser	Lys	Met	Arg	Tyr	Phe	Ala	Glu	Asn	Arg	Leu	Tyr	Ser	225	230	235
Cys	Gln	Phe	Asp	Pro	Glu	Ser	Phe	Lys	Thr	Ile	Ser	Ala	Val	Glu	Arg	245	250	255
Asp	Arg	Asn	Gly	Tyr	Tyr	Ile	Lys	Arg	Lys	Phe	Gln	Glu	Gln	Gln	Arg	260	265	270
Ile	Ala	Ser	Asn	Phe	Lys	Lys	Cys	Thr	Ile	Tyr	Arg	Lys	Met	Thr	Ser	275	280	285
Phe	Met	Leu	Gln	Thr	Leu	Glu	Ser	Leu	Val	Val	His	Lys	Ile	Tyr	Leu	290	295	300
Asn	Ser	Ile	Thr	Tyr	Glu	Ile	Lys	Arg	Gln	Val	Asn	His	Lys	Ile	His	305	310	315
Gly	Leu	Ile	Tyr	Gln	Lys	Ile	Ile	His	Thr	Ala	Asn	Ser	Leu	Leu	Trp	325	330	335
Leu	Val	Leu	Thr	Asn	Ser	Glu	Ile	Leu	Phe	Gly	Asp	Asn	Val	Lys	Gly	340	345	350
Leu	Asn	Trp	Leu	Arg	Lys	Lys	Glu	Ser	Leu	Lys	Val	Asp	Arg	Ser	Ile	355	360	365
Ile	Lys	Ile	Thr	Gln	Glu	Ile	Met	Arg	Arg	Lys	Leu	Tyr	Lys	Glu	Gly	370	375	380
Asn	Met	Thr	Val	Asn	Gln	Ile	Cys	Glu	Ile	Thr	Asn	Val	Ser	Arg	Ala	385	390	395
Ser	Leu	Tyr	Arg	Lys	Leu	Ser	Glu	Val	Asn	Asn	Pro	Phe	Cys	Ile	Pro	405	410	415
Leu	Met	Gly	Asn	Ile	Phe	Lys	Glu	Glu	Lys	Glu	Thr	Ile	Lys	Tyr	Gln	420	425	430
Pro	Pro	Ser	Asp	Ala	Glu	Lys	Pro	Phe	Asp	Lys	Lys	Arg	Ile	Ile	Ile	435	440	445
Leu	Arg	Asn	Ser	Ser	Phe	Ile	Met	Met	Leu	Ile	Asn	Ser	Ala	Leu	Ser	450	455	460

Asp	Lys	Leu	Leu	Arg	Ala	Asn	Leu	Cys	Glu	Arg	Val	Ile	Thr	Met	Ser	
465					470					475					480	
Asp	Lys	Ile	Leu	Ile	Val	Asp	Asp	Glu	His	Glu	Ile	Ala	Asp	Leu	Val	
				485					490					495		
Glu	Leu	Tyr	Leu	Lys	Asn	Glu	Asn	Tyr	Thr	Val	Phe	Lys	Tyr	Tyr	Thr	
			500					505					510			
Ala	Lys	Glu	Ala	Leu	Glu	Cys	Ile	Asp	Lys	Ser	Glu	Ile	Asp	Leu	Ala	
		515					520					525				
Ile	Leu	Asp	Ile	Met	Leu	Pro	Gly	Thr	Ser	Gly	Leu	Thr	Ile	Cys	Gln	
	530					535					540					
Lys	Ile	Arg	Asp	Lys	His	Thr	Tyr	Pro	Ile	Ile	Met	Leu	Thr	Gly	Lys	
545					550					555					560	
Asp	Thr	Glu	Val	Asp	Lys	Ile	Thr	Gly	Leu	Thr	Ile	Gly	Ala	Asp	Asp	
				565					570					575		
Tyr	Ile	Thr	Lys	Pro	Phe	Arg	Pro	Leu	Glu	Leu	Ile	Ala	Arg	Val	Lys	
			580					585					590			
Ala	Gln	Leu	Arg	Arg	Tyr	Lys	Lys	Phe	Ser	Gly	Val	Lys	Glu	Gln	Asn	
		595					600					605				
Glu	Asn	Val	Ile	Val	His	Ser	Gly	Leu	Val	Ile	Asn	Val	Asn	Thr	His	
	610					615					620					
Glu	Cys	Tyr	Leu	Asn	Glu	Lys	Gln	Leu	Ser	Leu	Thr	Pro	Thr	Glu	Phe	
625					630					635					640	
Ser	Ile	Leu	Arg	Ile	Leu	Cys	Glu	Asn	Lys	Gly	Asn	Val	Val	Ser	Ser	
				645					650					655		
Glu	Leu	Leu	Phe	His	Glu	Ile	Trp	Gly	Asp	Glu	Tyr	Phe	Ser	Lys	Ser	
			660					665					670			
Asn	Asn	Thr	Ile	Thr	Val	His	Ile	Arg	His	Leu	Arg	Glu	Lys	Met	Asn	
		675					680					685				
Asp	Thr	Ile	Asp	Asn	Pro	Lys	Tyr	Ile	Lys	Thr	Val	Trp	Gly	Val	Gly	
						695					700					
Tyr	Lys	Ile	Glu	Lys	Lys	Lys	Arg	Leu	Phe	Gln	Thr	Arg	Thr	Lys	Thr	
705					710					715					720	
Leu	His	Val	Tyr	Arg	Cys	Asn	Cys	Cys	Gly	Ser	Asn	Cys	Ile	Arg	Val	
				725					730					735		
Val	Tyr	Ser	Phe	Asn	Asp	Pro	Arg	Glu	Thr	Trp	Gly	Leu	Asp	Leu	Lys	
				740				745					750			
Tyr	Phe	Gly	Lys	Gln	Ile	Leu	Lys	Ser	Pro	Gly	Arg	Asp	Glu	Ile	Ile	
		755					760					765				

Ser	Ile	Phe	His	Thr	Glu	Gln	Tyr	Arg	Tyr	Leu	Tyr	Leu	Cys	Gly	Asp	770	775	780	
Cys	His	Tyr	Ser	Tyr	Ser	Met	Ser	Arg	His	Ala	Phe	Lys	Ile	Arg	Lys	785	790	795	800
Ile	Leu	Arg	Asp	Lys	Tyr	Arg	His	Cys	Thr	Tyr	Ser	Glu	Arg	Arg	Thr	805	810	815	
Asn	Ala	Phe	Cys	Gly	Asn	Gly	Cys	Tyr	Gly	Thr	Lys	Ala	Gln	His	Ile	820	825	830	
Lys	Thr	Asp	Ser	Gly	Lys	Ala	Arg	Ala	Gly	Cys	Lys	Ala	Gly	Arg	Thr	835	840	845	
Lys	Lys	Lys	Arg	Cys	Tyr	Val	Leu	Gly	Ala	Arg	Tyr	Asn	Ala	Pro	Tyr	850	855	860	
Ile	His	Tyr	Arg	Leu	Phe	Glu	Pro	Ala	Arg	Gly	Ser	Arg	His	Ala	Gly	865	870	875	880
Arg	Ser	Lys	Gly	Lys	Val	Cys	Ala	Tyr	His	Val	Gly	Gln	Ser	Val	Ser	885	890	895	
Thr	Arg	Thr	Ala	Asn	Arg	Arg	Val	Phe	Asp	Tyr	Thr	Val	Pro	Thr	Asn	900	905	910	
Asp	Asn	Ala	Asn	Lys	Asn	Ala	His	Arg	Pro	Ile	Leu	Tyr	Ala	Gly	Ala	915	920	925	
Asp	Asp	Arg	Ile	Leu	Ser	Ser	Ala	Phe	Arg	Thr	Trp	Lys	Thr	Gly	Gly	930	935	940	
Tyr	Ser	Arg	Pro	Arg	Gly	Ser	Asp	Arg	Val	Arg	Arg	Pro	Thr	Arg	Glu	945	950	955	960
Ser	Leu	Gln	His	Phe	Glu	Lys	Arg	Arg	Cys	Ile	Gln	Gly	Gln	His	His	965	970	975	
His	Tyr	Arg	Gly	Pro	Leu	Arg	Gly	Cys	Gly	Val	Asn	Arg	Ile	Gln	Glu	980	985	990	
His	Trp	Lys	His	Pro	Lys	Arg	Ala	Ser	Cys	His	Ile	Lys	Val	Leu	Ala	995	1000	1005	
Gly	Gln	Phe	Ser	Phe	Phe	Arg	Tyr	Gly	Trp	Arg	Gly	Thr	Trp	Ile	Gly	1010	1015	1020	
Asp	Cys	Lys	Arg	Asn	Tyr	Cys	Ser	Ala	Trp	Arg	Ala	Asp	Leu	Arg	Gly	1025	1030	1035	1040
Lys	Leu	Leu	Tyr	Asp	Val	Gly	Arg	Ala	Ser	Ser	Asp	Ala	Arg	Leu	Gly	1045	1050	1055	
Lys	Glu	Val	Leu	Arg	Asp	Val	Tyr	Asn	Phe	Leu	Gly	Lys	Ser	Gln	Gly	1060	1065	1070	

Tyr	Leu	Tyr	Phe	Phe	Leu	Gly	Asn	Gln	Phe	Asn	Ile	Lys	Lys	Arg	Leu
	1075						1080					1085			
Val	Leu	Thr	Arg	Thr	Tyr	Arg	Lys	Asn	Glu	Pro	Phe	Ser	Phe	Phe	Arg
	1090					1095					1100				
Glu	Arg	Phe	Asp	Lys	Ile	Thr	Ile	Gly	Ile	Pro	Val	Leu	Phe	Gly	Ala
	1105				1110					1115					1120
Phe	His	Arg	Lys	Gly	Trp	Ser	Leu	Ile	Thr	Ser	Ala	Leu	Leu	Phe	Met
				1125					1130					1135	
Asp	Val	Ser	Arg	Met	Arg	Gln	Met	His	Ser	Met	Leu	Phe	Arg	Leu	Ala
			1140					1145					1150		
Leu	Ala	Leu	Trp	Gln	Arg	Leu	Thr	Pro	Thr	Cys	Arg	Asn	Pro	Thr	Pro
		1155					1160					1165			
Asn	Pro	Arg	Leu	Ser	Ile	Asn	Val	Ser	Val	Trp	Asp	Ile	Asn	Gln	Arg
	1170					1175					1180				
Phe	Pro	Pro	Leu	Phe	Phe	Leu	Arg	Arg	Glu	Pro	Val	Asn	Ile	Phe	Leu
	1185				1190					1195					1200
Pro	Glu	Ala	Ser	Ala	Ala	Ile	Ile	Ile	Gln	Leu	Leu	Leu	Arg	Glu	Trp
				1205					1210					1215	
Ala	Ser	Leu	Ser	Thr	Met	Trp	Arg	Thr	Arg	Arg	Ile	Ala	Leu	Pro	Ile
			1220					1225					1230		
Ile	Leu	Cys	Phe	Leu	Trp	Gln	Tyr	Ala	Thr	Asn	Arg	Leu	Cys	Ala	Leu
		1235					1240					1245			
Trp	Lys	Asn	Met	Ile	Ser	Gly	Trp	Thr	Ala	Thr	Val	Ala	Arg	Tyr	Ser
	1250					1255					1260				
Ala	Thr	Gln	Leu	Val	Trp	Trp	Glu	Arg	Ala	Arg	Ala	Lys	Arg	Leu	Leu
	1265				1270					1275					1280
Ser	Gly	Cys	Glu	Asp	Leu	Asp	Val	Lys	Cys	Trp	Leu	Ile	Val	Ala	Ala
				1285					1290					1295	
Glu	Val	Arg	Thr	Met	Tyr	Arg	Leu	Met	Ser	Cys	Cys	Lys	Ile	Ala	Ile
			1300					1305					1310		
Ser	Leu	Arg	Phe	Met	Cys	Arg	Ser	Ile	Arg	Ile	Arg	Thr	Ile	Leu	Ser
		1315					1320					1325			
Ala	Thr	Asn	Lys	Tyr	Arg	Glu	Ser	Lys	Glu	His	Phe	Leu	Ser	Ile	Leu
	1330					1335					1340				
Gly	Ala	Val	His	Leu	Ile	Pro	Met	Ser	Trp	Leu	Lys	His	Lys	Thr	Gly
	1345				1350					1355					1360
Asn	Trp	Ala	Val	Pro	His	Trp	Met	Tyr	Trp	Lys	Glu	Arg	Lys	Ser	Phe
				1365					1370					1375	

Ser Thr Leu Ile Ala Pro Lys Asn Gln Leu Ile Ile Asn Phe Tyr Leu
 1380 1385 1390
 Asn Phe Lys Glu Cys Leu Thr Ser His Arg Ile Arg Pro Ile Ile Pro
 1395 1400 1405
 Ser Lys Arg Cys Val Ile Pro Leu Lys Lys Pro Leu Lys Thr Val Trp
 1410 1415 1420
 Ile Leu Lys Gly Asp Arg Ser Met Asn Arg Ile Lys Val Ala Ile Leu
 1425 1430 1435 1440
 Phe Gly Gly Cys Ser Glu Glu His Asp Val Ser Val Lys Ser Ala Ile
 1445 1450 1455
 Glu Ile Ala Ala Asn Ile Asn Lys Glu Lys Tyr Glu Pro Leu Tyr Ile
 1460 1465 1470
 Gly Ile Thr Lys Ser Gly Val Trp Lys Met Cys Glu Lys Pro Cys Ala
 1475 1480 1485
 Glu Trp Glu Asn Asp Asn Cys Tyr Ser Ala Val Leu Ser Pro Asp Lys
 1490 1495 1500
 Lys Met His Gly Leu Leu Val Lys Lys Asn His Glu Tyr Glu Ile Asn
 1505 1510 1515 1520
 His Val Asp Val Ala Phe Ser Ala Leu His Gly Lys Ser Gly Glu Asp
 1525 1530 1535
 Gly Ser Ile Gln Gly Leu Phe Glu Leu Ser Gly Ile Pro Phe Val Gly
 1540 1545 1550
 Cys Asp Ile Gln Ser Ser Ala Ile Cys Met Asp Lys Ser Leu Thr Tyr
 1555 1560 1565
 Ile Val Ala Lys Asn Ala Gly Ile Ala Thr Pro Ala Phe Trp Val Ile
 1570 1575 1580
 Asn Lys Asp Asp Arg Pro Val Ala Ala Thr Phe Thr Tyr Pro Val Phe
 1585 1590 1595 1600
 Val Lys Pro Ala Arg Ser Gly Ser Ser Phe Gly Val Lys Lys Val Asn
 1605 1610 1615
 Ser Ala Asp Glu Leu Asp Tyr Ala Ile Glu Ser Ala Arg Gln Tyr Asp
 1620 1625 1630
 Ser Lys Ile Leu Ile Glu Gln Ala Val Ser Gly Cys Glu Val Gly Cys
 1635 1640 1645
 Ala Val Leu Gly Asn Ser Ala Ala Leu Val Val Gly Glu Val Asp Gln
 1650 1655 1660
 Ile Arg Leu Gln Tyr Gly Ile Phe Arg Ile His Gln Glu Val Glu Pro
 1665 1670 1675 1680

Glu Lys Gly Ser	Glu Asn Ala Val	Ile Thr Val Pro	Ala Asp Leu Ser
1685		1690	1695
Ala Glu Glu Arg	Gly Arg Ile Gln	Glu Thr Ala Lys	Lys Ile Tyr Lys
1700		1705	1710
Ala Leu Gly Cys	Arg Gly Leu Ala	Arg Val Asp Met	Phe Leu Gln Asp
1715		1720	1725
Asn Gly Arg Ile	Val Leu Asn Glu	Val Asn Thr Leu	Pro Gly Phe Thr
1730		1735	1740
Ser Tyr Ser Arg	Tyr Pro Arg Met	Met Ala Ala Ala	Gly Ile Ala Leu
1745		1750	1755
Pro Glu Leu Ile	Asp Arg Leu Ile	Val Leu Ala Leu	Lys Gly Ala Trp
	1765		1770
Lys Asp Leu Leu	Phe Met Lys Tyr	Thr Val Phe Val	Gly Thr Leu Asn
	1780		1785
Met Pro Leu Gly	Ile Ile Ser Pro	Glu Asn Arg Leu	Thr Val Met Lys
	1795		1800
Ile Ala Leu Gly	His Thr Ser Trp	Leu Asn Arg Phe	Arg Gln Lys Asn
	1810		1815
Trp Leu Leu Pro	Lys Gly Thr Asp	Cys Phe Tyr Gly	Thr Val Thr Val
	1825		1830
Leu Ser Val Leu	Thr Val Leu Cys	Asn Gly Leu His	Ser Arg Lys Ile
	1845		1850
Thr Gln Arg Lys	Val Ile Ile Pro	Ile Leu Thr Glu	Leu Arg Phe Gln
	1860		1865
Lys Asp Thr Trp	Leu Gln Asn Gln	Ala Ile Ala Ala	Ala Val Pro Leu
	1875		1880
Ile Leu Arg Phe	Ile Asp Thr Arg	Val Ser Leu Tyr	Gln Trp Gly Ala
	1890		1895
Asp Leu Ile Leu	Trp Met Asn Ala	Leu Ile Met Arg	Gln Met Glu Tyr
	1905		1910
His Ala Met Lys	Arg Lys Ile Ala	Asp Val Cys Ala	Pro Ser Trp Lys
	1925		1930
Thr Val Gly Leu	Lys His Ile Ala	Ser Asn Gly Gly	Thr Met Tyr Glu
	1940		1945
Thr Asn His Thr	Pro Ile Ala Ile	Leu Ile Ser Pro	Leu Asn Lys Leu
	1955		1960
Leu Thr Val Ala	Arg Thr Asn Tyr	Ile Ser Leu Phe	Arg Gln Glu Thr
	1970		1975
			1980

Arg Arg Met Leu Val Leu Arg Glu Phe Ile Tyr Ser Arg Tyr Arg Cys
 1985 1990 1995 2000
 Lys Ala Glu Arg Tyr Cys Gly His Tyr Leu Arg Ala Leu Arg Gln Asp
 2005 2010 2015
 Ser Leu Ile Ile Arg Leu Ile Ala Arg Gly Gly Ile Ser His Arg Pro
 2020 2025 2030
 Leu Ser Thr Gly Ser Ser Ala Ser Leu Asn Ser Ala Trp Val Ser Leu
 2035 2040 2045
 Met Lys Ile His Leu His Trp Ile Gln Gly Glu Ile Ile Asp Cys Asn
 2050 2055 2060
 Leu Arg Gly Lys Thr Ala Gln Ser Gln Thr Arg Leu Cys Arg Leu Arg
 2065 2070 2075 2080
 Gly Arg Phe Lys Tyr Phe Ile Leu Pro Thr Ile Leu Arg Arg Arg Leu
 2085 2090 2095
 Lys Met Lys Lys Leu Phe Phe Leu Leu Leu Leu Phe Leu Ile Tyr
 2100 2105 2110
 Leu Gly Tyr Asp Tyr Val Asn Glu Ala Leu Phe Ser Gln Glu Lys Val
 2115 2120 2125
 Glu Phe Gln Asn Tyr Asp Gln Asn Pro Lys Glu His Leu Glu Asn Ser
 2130 2135 2140
 Gly Thr Ser Glu Asn Thr Gln Glu Lys Thr Ile Thr Glu Glu Gln Val
 2145 2150 2155 2160
 Tyr Gln Gly Asn Leu Leu Leu Ile Asn Ser Lys Tyr Pro Val Arg Gln
 2165 2170 2175
 Glu Val Ser Gln Ile Ser Ile Tyr Leu Asn Met Thr Asn Met Asp Thr
 2180 2185 2190
 Gly Cys Leu Ile Val Ile Phe Ile Cys Gln Lys Lys His Lys Asn Phe
 2195 2200 2205
 Gln Arg Trp Ser Met Met Leu Arg Val Ala Leu Val Ile Leu Leu Leu
 2210 2215 2220
 Ile Val Ala Ile Glu Thr Leu Met Ser Lys Val Cys Phe Thr Lys Lys
 2225 2230 2235 2240
 Trp Gly Leu Ser Met Pro Tyr Gln Gln Val Ile Val Ser Ile Ile Gln
 2245 2250 2255
 Val Tyr His Met Asp Gln Ala Arg Lys Trp Asn Glu Pro Leu Lys Glu
 2260 2265 2270
 Ser Gly Lys Lys Met Leu Gly Asn Thr Gly Ser Phe Tyr Val Ile Gln
 2275 2280 2285

Arg Thr Lys Gln Ser Gln Glu Phe
2290 2295

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2254 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ser	Phe	Ser	Phe	Cys	Ser	Phe	Val	Arg	Asp	Leu	Leu	Thr	Val	Leu	Asn	
1				5				10						15		
Ser	Phe	Phe	Ser	Ala	Ile	Ala	Leu	Ala	Ser	His	Thr	Ile	Leu	Ser	Ser	
			20					25					30			
Val	Val	Ile	Ala	Gly	Ser	Ile	Ile	Leu	Phe	Phe	Leu	Arg	Lys	Ser	Met	
		35					40					45				
His	Ser	Cys	Ser	Arg	Met	Ala	Ser	Pro	Phe	Ser	Lys	Ala	Asn	Arg	Tyr	
	50					55					60					
Leu	Asn	Val	Ile	Arg	Tyr	Ser	Leu	Arg	Val	Lys	Val	Thr	Lys	Ser	Tyr	
65					70					75					80	
Ser	Leu	Arg	Ile	Ser	Phe	Lys	Ser	Gln	Ser	Val	Phe	Ser	Leu	Gly	Ser	
			85					90						95		
Ser	Glu	Asp	Gly	Leu	Thr	Pro	Ile	Cys	Phe	Asp	Ile	Tyr	Cys	Met	Thr	
			100					105					110			
Glu	Ser	Gly	Met	Leu	Leu	Ile	Val	Tyr	Gly	Gln	Pro	Gly	Tyr	Arg	Arg	
		115					120					125				
Thr	Ala	Asn	Thr	Ala	Asn	Pro	Lys	Arg	Phe	Ser	Ser	Leu	Leu	Arg	Leu	
						135						140				
Leu	Thr	Ile	Ser	Lys	Ser	Arg	Leu	Glu	Lys	Val	Lys	Val	Pro	Ser	Ile	
145					150					155					160	
His	Ser	Ser	Ser	Gly	Ile	Cys	Ile	Lys	Ala	Cys	Leu	Cys	Ser	Gly	Val	
				165					170					175		
Ser	Asn	Ser	Leu	Pro	Leu	Ala	Ile	Phe	Ile	Gln	Tyr	His	Ser	Ile	Ser	
			180					185					190			
Val	Phe	Ser	Ile	Tyr	Phe	Asn	Tyr	Ile	Ser	Ile	Glu	Cys	Thr	Leu	Leu	
		195					200					205				
Ile	Gln	Met	Thr	Asp	Lys	Ile	Ile	Val	Lys	Ser	Val	Ser	Asp	Leu	Ser	
		210				215						220				

Gln Lys Gly Asp Ile Leu Arg Lys Ile Gly Tyr Ile Arg Val Ser Ser
 225 230 235 240
 Thr Asn Gln Asn Pro Ser Arg Gln Phe Gln Gln Leu Asn Glu Ile Gly
 245 250 255
 Met Asp Ile Ile Arg Glu Ser Phe Arg Ser Asn Lys Gly Ser Arg Ala
 260 265 270
 Thr Ser Lys Ser Val Arg Arg Phe Thr Gly Arg His His Leu Cys Tyr
 275 280 285
 Arg Leu Asn Ser Asn His Ser Tyr Thr Arg Ser Ile Ile Asn Arg His
 290 295 300
 Thr Arg Lys Gly Lys Phe Lys Ile Thr Lys Arg Tyr Met Ala Phe Ile
 305 310 315 320
 Arg Arg Ser Ile Gln Pro Ile Leu Asn Tyr Cys Asn Gly Trp Cys Pro
 325 330 335
 Ile Arg Ala Arg Ser Tyr Ser Asp Glu Thr Thr Arg Asp Ile Gly Glu
 340 345 350
 Arg Arg Lys Val Arg Ser Ile Lys Glu Val Ser Lys Ser Arg Arg Asn
 355 360 365
 Glu Leu Cys Gly Glu Ser Tyr Ile Lys Lys Glu Ile Leu Ile Lys Phe
 370 375 380
 Val Lys Leu Leu Met Tyr Leu Gly Leu His Tyr Thr Gly Asn Tyr Gln
 385 390 395 400
 Lys Ile Ile Ser His Ser Val Phe Arg Trp Ala Ile Phe Leu Lys Lys
 405 410 415
 Lys Arg Lys Leu Asn Ile Asn Ser Leu Leu Ala Met Pro Lys Ser Pro
 420 425 430
 Leu Ile Lys Lys Glu Ser Ser Ser Glu Ile Leu Ser His Leu Leu Cys
 435 440 445
 Lys Cys Leu Ile Arg Pro Tyr Asn Leu Ile Asn Tyr Gly Gln Thr Tyr
 450 455 460
 Val Lys Gly Leu Ala Ile Lys Tyr Leu Leu Trp Met Met Asn Met Lys
 465 470 475 480
 Leu Pro Ile Trp Leu Asn Tyr Thr Lys Thr Arg Ile Ile Arg Phe Ser
 485 490 495
 Asn Thr Ile Pro Pro Lys Lys His Trp Asn Val Thr Ser Leu Arg Leu
 500 505 510
 Thr Leu Pro Tyr Trp Thr Ser Cys Phe Pro Ala Gln Ala Ala Leu Leu
 515 520 525

Ser	Val	Lys	Lys	Gly	Thr	Ser	Thr	Pro	Ile	Arg	Leu	Ser	Cys	Pro	Gly	530	535	540	
Lys	Ile	Gln	Arg	Ile	Lys	Leu	Gln	Gly	Gln	Ser	Ala	Arg	Met	Ile	Ile	545	550	555	560
Arg	Ser	Pro	Phe	Ala	His	Trp	Ser	Leu	Leu	Gly	Arg	Pro	Ser	Cys	Ala	565	570	575	
Asp	Thr	Lys	Asn	Ser	Val	Glu	Arg	Ser	Arg	Thr	Lys	Met	Leu	Ser	Ser	580	585	590	
Thr	Pro	Ala	Leu	Ser	Leu	Met	Leu	Thr	Pro	Met	Ser	Val	Ile	Thr	Arg	595	600	605	
Ser	Ser	Tyr	Pro	Leu	Leu	Pro	Pro	Ser	Phe	Gln	Tyr	Cys	Glu	Ser	Ser	610	615	620	
Val	Lys	Thr	Arg	Gly	Met	Trp	Leu	Ala	Pro	Ser	Cys	Tyr	Phe	Met	Arg	625	630	635	640
Tyr	Gly	Ala	Thr	Asn	Ile	Ser	Ala	Arg	Ala	Thr	Thr	Pro	Ser	Pro	Cys	645	650	655	
Ile	Ser	Gly	Ile	Cys	Ala	Lys	Lys	Thr	Thr	Pro	Leu	Ile	Ile	Arg	Asn	660	665	670	
Ile	Lys	Arg	Tyr	Gly	Gly	Leu	Val	Ile	Lys	Leu	Lys	Asn	Lys	Lys	Asn	675	680	685	
Asp	Tyr	Ser	Lys	Leu	Glu	Arg	Lys	Leu	Tyr	Met	Tyr	Ile	Val	Ala	Ile	690	695	700	
Val	Val	Val	Ala	Ile	Val	Phe	Val	Leu	Tyr	Ile	Arg	Ser	Met	Ile	Arg	705	710	715	720
Gly	Lys	Leu	Gly	Asp	Trp	Ile	Leu	Ser	Ile	Leu	Glu	Asn	Lys	Tyr	Asp	725	730	735	
Leu	Asn	His	Leu	Asp	Ala	Met	Lys	Leu	Tyr	Gln	Tyr	Ser	Ile	Arg	Asn	740	745	750	
Asn	Ile	Asp	Ile	Phe	Ile	Tyr	Val	Ala	Ile	Val	Ile	Ser	Ile	Leu	Ile	755	760	765	
Leu	Cys	Arg	Val	Met	Leu	Ser	Lys	Phe	Ala	Lys	Tyr	Phe	Asp	Glu	Ile	770	775	780	
Asn	Thr	Gly	Ile	Asp	Val	Leu	Ile	Gln	Asn	Glu	Asp	Lys	Gln	Ile	Glu	785	790	795	800
Leu	Ser	Ala	Glu	Met	Asp	Val	Met	Glu	Gln	Lys	Leu	Asn	Thr	Leu	Lys	805	810	815	
Arg	Thr	Leu	Glu	Lys	Arg	Glu	Gln	Asp	Ala	Lys	Leu	Ala	Glu	Gln	Arg	820	825	830	

Lys	Asn	Asp	Val	Val	Met	Tyr	Leu	Ala	His	Asp	Ile	Lys	Thr	Pro	Leu		
		835					840					845					
Thr	Ser	Ile	Ile	Gly	Tyr	Leu	Ser	Leu	Leu	Asp	Glu	Ala	Pro	Asp	Met		
	850					855					860						
Pro	Val	Asp	Gln	Lys	Ala	Lys	Tyr	Val	His	Ile	Thr	Leu	Asp	Lys	Ala		
865					870					875					880		
Tyr	Arg	Leu	Glu	Gln	Leu	Ile	Asp	Glu	Phe	Phe	Glu	Ile	Thr	Arg	Tyr		
				885					890					895			
Asn	Leu	Gln	Thr	Ile	Thr	Leu	Thr	Lys	Thr	His	Ile	Asp	Leu	Tyr	Tyr		
			900					905					910				
Met	Leu	Val	Gln	Met	Thr	Asp	Glu	Phe	Tyr	Pro	Gln	Leu	Ser	Ala	His		
		915					920					925					
Gly	Lys	Gln	Ala	Val	Ile	His	Ala	Pro	Glu	Asp	Leu	Thr	Val	Ser	Gly		
	930					935					940						
Asp	Pro	Asp	Lys	Leu	Ala	Arg	Val	Phe	Asn	Asn	Ile	Leu	Lys	Asn	Ala		
945					950					955					960		
Ala	Ala	Tyr	Ser	Glu	Asp	Asn	Ser	Ile	Ile	Asp	Ile	Thr	Ala	Gly	Leu		
				965					970					975			
Ser	Gly	Asp	Val	Val	Ser	Ile	Glu	Phe	Lys	Asn	Thr	Gly	Ser	Ile	Pro		
			980					985					990				
Lys	Asp	Lys	Leu	Ala	Ala	Ile	Phe	Glu	Lys	Phe	Tyr	Arg	Leu	Asp	Asn		
		995					1000					1005					
Ser	Arg	Ser	Ser	Asp	Thr	Gly	Gly	Ala	Gly	Leu	Gly	Leu	Ala	Ile	Ala		
	1010					1015					1020						
Lys	Glu	Ile	Ile	Val	Gln	His	Gly	Gly	Gln	Ile	Tyr	Ala	Glu	Ser	Tyr		
1025					1030					1035					1040		
Asp	Asn	Tyr	Thr	Thr	Phe	Arg	Val	Glu	Leu	Pro	Ala	Met	Pro	Asp	Leu		
				1045					1050					1055			
Val	Asp	Lys	Arg	Arg	Ser	Glu	Met	Tyr	Ile	Ile	Phe	Glu	Asn	Leu	Lys		
			1060					1065					1070				
Val	Ile	Phe	Thr	Phe	Ser	Glu	Ile	Asn	Asn	Leu	Ile	Leu	Arg	Asn	Gly		
	1075						1080					1085					
Ser	Phe	Leu	His	Gly	Arg	Leu	Asn	Thr	Val	Arg	Thr	Ser	Arg	Phe	Arg		
	1090					1095					1100						
Ser	Ser	Glu	Lys	Asp	Leu	Thr	Arg	Leu	Pro	Leu	Ala	Ser	Pro	Phe	Tyr		
1105					1110					1115					1120		
Leu	Val	Pro	Phe	Thr	Glu	Arg	Val	Gly	Leu	Asn	Tyr	Glu	His	Arg	His		
				1125					1130					1135			

Tyr Cys Leu Trp Met Ala Gly Gly Arg Cys Ile Pro Cys Ser Phe Ala
 1140 1145 1150
 Ser Leu Trp Arg Tyr Gly Asn Asp Asn Arg Gln Arg Val Gly Ile Gln
 1155 1160 1165
 Arg Gln Ile Arg Ala Phe Gln Ser Met Tyr Gln Cys Gly Thr Ile Arg
 1170 1175 1180
 Asp Phe Arg Leu Tyr Ser Ser Cys Ala Glu Glu Ser Arg Cys Glu Ile
 1185 1190 1195 1200
 Tyr Phe Tyr Pro Lys His Arg Leu Gln Ser Tyr Arg Tyr Asn Cys Cys
 1205 1210 1215
 Glu Asn Gly His His Cys Arg Gln Cys Gly Val Leu Ala Gly Arg Cys
 1220 1225 1230
 Arg Leu Tyr Tyr Asp Ala Asn Ser Tyr Gly Ser Thr Gln Arg Lys Ile
 1235 1240 1245
 Asp Cys Ala Leu Cys Gly Lys Thr Phe Gln Val Gly Gln Arg Pro Trp
 1250 1255 1260
 Gln Gly Thr Gln Arg His Asp Ser Trp Cys Gly Gly Asn Gly Pro Asp
 1265 1270 1275 1280
 Arg Gln Ser Gly Tyr Ala Ala Ala Arg Ile Trp Met Ser Val Gly Leu
 1285 1290 1295
 Ser Gln Pro Lys Tyr Arg Gly Lys Leu Cys Thr Val Val Ala Ala Lys
 1300 1305 1310
 Arg Tyr Arg Tyr Ala Ser Cys Ala Ala Gln Tyr Gly Tyr Ala Leu Tyr
 1315 1320 1325
 Tyr Gln Pro Arg Thr Asn Thr Glu Asn Glu Ala Arg Ser Ile Ser Tyr
 1330 1335 1340
 Gln Tyr Trp Ala Arg Ser Thr Cys Arg Tyr Leu Val Gly Ser Ile Arg
 1345 1350 1355 1360
 Lys Arg Glu Thr Gly Arg Cys Arg Ile Gly Cys Ile Gly Arg Arg Gly
 1365 1370 1375
 Arg Val Phe Leu Leu Leu His Pro Lys Thr Asn Ser Ile Phe Thr Thr
 1380 1385 1390
 Ser Lys Asn Ala Arg Asp Asn His Thr Ala Tyr Gly Leu Leu Tyr Arg
 1395 1400 1405
 Ala Ser Val Ala Tyr Arg Lys Asn His Lys Leu Phe Gly Phe Lys Glu
 1410 1415 1420
 Thr Gly Ala Ile Glu Lys Leu Gln Tyr Cys Leu Gly Val Ala Gln Arg
 1425 1430 1435 1440

Ser Met Thr Tyr Arg Asn Leu Gln Arg Pro Leu Thr Leu Ile Lys Lys
 1445 1450 1455
 Asn Thr Ser Arg Tyr Thr Leu Glu Leu Arg Asn Leu Val Tyr Gly Lys
 1460 1465 1470
 Cys Ala Lys Asn Leu Ala Arg Asn Gly Lys Thr Thr Ile Ala Ile Gln
 1475 1480 1485
 Leu Tyr Ser Arg Arg Ile Lys Lys Cys Thr Asp Tyr Leu Leu Lys Arg
 1490 1495 1500
 Thr Met Asn Met Lys Ser Thr Met Leu Met His Phe Gln Leu Cys Met
 1505 1510 1515 1520
 Ala Ser Gln Val Lys Met Asp Pro Tyr Lys Val Cys Leu Asn Cys Pro
 1525 1530 1535
 Val Ser Leu Leu Ala Ala Ile Phe Lys Ala Gln Gln Phe Val Trp Thr
 1540 1545 1550
 Asn Arg His Thr Ser Leu Arg Lys Met Leu Gly Leu Leu Pro Pro Phe
 1555 1560 1565
 Gly Leu Leu Ile Lys Met Ile Gly Arg Trp Gln Leu Arg Leu Pro Ile
 1570 1575 1580
 Leu Phe Leu Leu Ser Arg Arg Val Gln Ala His Pro Ser Val Lys Lys
 1585 1590 1595 1600
 Ser Ile Ala Arg Thr Asn Trp Thr Thr Gln Leu Asn Arg Gln Asp Asn
 1605 1610 1615
 Met Thr Ala Lys Ser Leu Ser Arg Leu Phe Arg Ala Val Arg Ser Val
 1620 1625 1630
 Val Arg Tyr Trp Glu Thr Val Pro Arg Leu Leu Ala Arg Trp Thr Lys
 1635 1640 1645
 Ser Gly Cys Ser Thr Glu Ser Phe Val Phe Ile Arg Lys Ser Ser Arg
 1650 1655 1660
 Lys Lys Ala Leu Lys Thr Gln Leu Pro Phe Pro Gln Thr Phe Gln Gln
 1665 1670 1675 1680
 Arg Ser Glu Asp Gly Tyr Arg Lys Arg Gln Lys Lys Tyr Ile Lys Arg
 1685 1690 1695
 Ser Ala Val Glu Val Pro Val Trp Ile Cys Phe Tyr Lys Ile Thr Ala
 1700 1705 1710
 Ala Leu Tyr Thr Lys Ser Ile Leu Cys Pro Val Ser Arg His Thr Val
 1715 1720 1725
 Val Ile Pro Val Trp Pro Leu Gln Val Leu His Phe Pro Asn Leu Thr
 1730 1735 1740

Ala Ser Tyr Arg Arg Gly Asp Lys His Gly Asn Arg Ile Tyr Phe Phe 1745	1750	1755	1760
Arg Asn Ser Thr Arg Cys Ser Leu Gly Arg Ile Cys His Leu Gly Phe 1765	1770	1775	
His Arg Lys Thr Gly Arg Leu Ser Lys Ser His Cys Arg Asp Ile Arg 1780	1785	1790	
Val Gly Ile Ala Phe Glu Gly Lys Arg Thr Gly Cys Tyr Pro Arg Val 1795	1800	1805	
Arg Ile Ala Ser Met Gly Arg Leu Pro Ser Ala Cys Cys Lys Leu Phe 1810	1815	1820	
Tyr Ala Met Gly Cys Thr Ala Gly Lys Pro Asp Lys Gly Lys Leu Leu 1825	1830	1835	1840
Ser Gln Tyr Pro Asn Asp Asp Phe Lys Arg Ile Arg Gly Phe Lys Ile 1845	1850	1855	
Lys Pro Pro Arg Gln Cys His Ser Tyr Ala Leu Ser Ile Arg His Gly 1860	1865	1870	
Ala Cys Thr Asn Gly Glu Pro Ile Phe Tyr Gly Thr Leu Ser Ser Cys 1875	1880	1885	
Gly Lys Trp Asn Ile Met Gln Ser Ala Lys Ser Gln Thr Phe Ala Leu 1890	1895	1900	
His His Gly Lys Gln Trp Val Ser Ile Pro Arg Met Val Ala Leu Cys 1905	1910	1915	1920
Ile Lys Arg Arg Thr Ile Pro Gln Leu Phe Phe Pro Arg Ile Asn Phe 1925	1930	1935	
Pro Leu His Gly Gln Thr Ile Ala Asn Ser Phe Gly Arg Lys Pro Asp 1940	1945	1950	
Val Cys Asn Trp Phe Leu Gly Asn Leu Tyr Ile Val Asp Ser Ile Glu 1955	1960	1965	
Asp Val Arg Gln Ser Asp Ile Ala Val Ile Ile Cys Val Arg Cys Gly 1970	1975	1980	
Lys Ile Ala Asp Ser His Arg Gly Val Val Phe His Thr Ala His Cys 1985	1990	1995	2000
Gln Gln Ala Val Gln Pro Arg Ile Gln His Gly Tyr His Leu Lys Phe 2005	2010	2015	
Ile Tyr Ile Gly Asp Asn Ser Lys Ser Ser Arg Ala Lys Leu Thr Val 2020	2025	2030	
Ile Tyr Gly Ala Lys Arg His Asn Leu Lys Arg Asp Cys Ala Val Gly 2035	2040	2045	

Glu 2050	Asp	Ser	Arg	Asn	Ile	Ser	Tyr	Phe	Gln	Leu	Tyr	Ser	Gly	Gly	Asp
Lys 2065	Arg	Ser	Cys	Phe	Phe	Tyr	Cys	Tyr	Cys	Tyr	Ser	Tyr	Thr	Val	Met 2080
Thr	Thr	Leu	Met	Lys 2085	His	Cys	Phe	Leu	Arg 2090	Lys	Lys	Ser	Asn	Phe 2095	Lys
Ile	Met	Ile	Lys 2100	Ile	Pro	Lys	Asn	Ile 2105	Lys	Ile	Val	Gly	Leu 2110	Leu	Lys
Ile	Pro	Lys 2115	Arg	Lys	Gln	Leu	Gln 2120	Lys	Asn	Arg	Phe	Ile 2125	Lys	Glu	Ile
Cys	Tyr 2130	Ser	Ile	Val	Asn	Ile 2135	Leu	Phe	Ala	Lys	Lys 2140	Cys	Glu	Val	Arg
Tyr 2145	Arg	Glu	Phe	Ile	Thr 2150	Arg	Ile	Asn	Lys	Trp 2155	Ile	Arg	Val	Ala	Tyr 2160
Leu	Tyr	Val	Lys 2165	Arg	Asn	Ser	Thr	Lys	Ile 2170	Phe	Arg	Asp	Gly	Gln 2175	Cys
Cys	Lys	Gly	Trp 2180	Arg	Ser	Phe	Tyr	Tyr 2185	Trp	Leu	Ser	Arg	Leu 2190	Ala	Lys
Cys	Ala 2195	Leu	Pro	Arg	Asn	Gly	Gly 2200	Val	Cys	Leu	Thr	Ser 2205	Arg	Leu	Ala
Phe 2210	Arg	Phe	Ile	Thr	Arg	Cys 2215	Arg	Ile	Lys	Leu	Asp 2220	Glu	Asn	Gly	Thr
Ser 2225	Pro	Arg	Lys	Val	Asp 2230	Arg	Arg	Lys	Cys	Leu 2235	Glu	Ile	Arg	Val	His 2240
Phe	Thr	Leu	Ser	Arg 2245	Gly	Gln	Asn	Arg	Val 2250	Asn	Arg	Asn	Ser		

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2291 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Ala Phe Leu Phe Ala His Leu Leu Glu Ile Tyr Pro Tyr Ile Ala Ser
1 5 10 15

Phe Gln Pro Leu Pro Leu Leu Pro Thr Pro Phe Phe Gln Val Gln Ala
20 25 30

Val	Phe	Cys	Phe	Phe	Leu	Glu	Asn	Leu	Cys	Ile	His	Ala	Val	Asp	Glu
		35					40					45			
Trp	His	His	His	Phe	Pro	Lys	Leu	Ile	Asp	Glu	Gly	Thr	Met	Ser	Phe
	50					55					60				
Asp	Ile	His	Ser	Gly	Lys	Leu	Gln	Ser	Arg	Ile	His	Phe	Glu	Phe	Leu
65					70					75					80
Ser	Asn	Asp	Pro	Lys	Val	Tyr	Phe	Pro	Phe	Glu	Asp	Asn	Asp	Gln	Ala
				85					90					95	
Arg	Met	Asp	His	Gln	Ser	Val	Ser	Ile	Tyr	Ile	Val	Pro	Asn	Leu	Gly
			100					105					110		
Cys	Phe	Tyr	Glu	Cys	Met	Ala	Asn	Arg	Asp	Thr	Glu	Glu	Gln	Leu	Ile
		115					120					125			
Glu	Gln	Gln	Ile	Leu	Asn	Gly	Phe	Leu	Pro	Ser	Phe	Ala	Tyr	Leu	Phe
	130					135					140				
Leu	Asn	Pro	Val	Trp	Lys	Lys	Ser	Arg	Ser	Pro	Val	Ser	Ile	His	Leu
145					150					155					160
Gln	Gly	Phe	Ala	Lys	Pro	Val	Ser	Val	Pro	Val	Ala	Ile	Leu	Tyr	Leu
				165					170					175	
Ser	Gln	Phe	Ser	Phe	Ser	Ile	Ile	Pro	Phe	Leu	Tyr	Phe	Gln	Phe	Ile
			180					185					190		
Ser	Ser	Ile	Ile	Tyr	Gln	Ser	Val	Leu	Tyr	Tyr	Lys	Cys	Ser	Arg	Leu
		195					200					205			
Ile	Lys	Ser	Leu	Arg	Ala	Ser	His	Lys	Thr	Cys	Leu	Lys	Asn	Glu	Val
	210					215					220				
Ile	Phe	Cys	Gly	Lys	Ser	Val	Ile	Phe	Val	Ser	Val	Arg	Leu	Thr	Arg
225					230					235					240
Ile	Leu	Gln	Asp	Asn	Phe	Ser	Ser	Thr	Arg	Ser	Glu	Trp	Ile	Leu	Tyr
				245					250					255	
Lys	Glu	Lys	Val	Ser	Gly	Ala	Thr	Lys	Asp	Arg	Glu	Gln	Leu	Gln	Lys
			260					265					270		
Val	Leu	Asp	Asp	Leu	Gln	Glu	Asp	Asp	Ile	Ile	Tyr	Val	Thr	Asp	Leu
		275					280					285			
Thr	Arg	Ile	Thr	Arg	Ser	Thr	Gln	Asp	Leu	Phe	Glu	Leu	Ile	Asp	Asn
	290					295					300				
Ile	Arg	Asp	Lys	Lys	Ala	Ser	Leu	Lys	Ser	Leu	Lys	Asp	Thr	Trp	Leu
305					310					315					320
Asp	Leu	Ser	Glu	Asp	Asn	Pro	Tyr	Ser	Gln	Phe	Leu	Ile	Thr	Val	Met
				325					330					335	

Ala	Gly	Val	Asn	Gln	Leu	Glu	Arg	Asp	Leu	Ile	Arg	Met	Arg	Gln	Arg		
			340					345					350				
Glu	Gly	Ile	Glu	Leu	Ala	Lys	Lys	Glu	Gly	Lys	Phe	Lys	Gly	Arg	Leu		
		355					360					365					
Lys	Lys	Tyr	His	Lys	Asn	His	Ala	Gly	Met	Asn	Tyr	Ala	Ala	Lys	Ala		
		370				375					380						
Ile	Arg	Arg	Lys	Tyr	Asp	Cys	Lys	Ser	Asn	Leu	Asn	Tyr	Cys	Ile	Gly		
385					390					395					400		
Phe	Ile	Ile	Gln	Glu	Ile	Ile	Arg	Ser	Glu	Leu	Ala	Ile	Leu	Tyr	Ser		
				405					410					415			
Ala	Asn	Gly	Gln	Tyr	Phe	Arg	Arg	Lys	Gly	Asn	Tyr	Lys	Ile	Leu	Thr		
			420					425					430				
Ala	Ser	Arg	Cys	Arg	Lys	Ala	Leu	Lys	Lys	Asn	His	His	Leu	Lys	Lys		
		435					440					445					
Phe	Leu	Val	Ile	Tyr	Tyr	Val	Asn	Ala	Tyr	Lys	Phe	Gly	Pro	Ile	Ile		
	450					455					460						
Ile	Ile	Lys	Gly	Lys	Leu	Met	Lys	Gly	Asp	Asn	Tyr	Glu	Arg	Asn	Thr		
465					470					475					480		
Tyr	Cys	Gly	Thr	Asn	Cys	Arg	Phe	Gly	Ile	Ile	Leu	Lys	Lys	Arg	Glu		
				485					490					495			
Leu	Tyr	Gly	Phe	Gln	Ile	Leu	Tyr	Arg	Gln	Arg	Ser	Ile	Gly	Met	Tyr		
			500					505					510				
Arg	Gln	Val	Asp	Pro	Cys	His	Ile	Gly	His	His	Ala	Ser	Arg	His	Lys		
		515					520					525					
Arg	Pro	Tyr	Tyr	Leu	Ser	Lys	Asn	Lys	Gly	Gln	Ala	His	Leu	Ser	Asp		
	530					535					540						
Tyr	His	Ala	Asp	Arg	Glu	Arg	Tyr	Arg	Gly	Arg	Asn	Tyr	Arg	Val	Asn		
545					550					555					560		
Asn	Arg	Arg	Gly	Leu	Tyr	Asn	Glu	Ala	Leu	Ser	Pro	Thr	Gly	Val	Asn		
				565					570					575			
Cys	Ser	Gly	Lys	Gly	Pro	Val	Ala	Pro	Ile	Gln	Lys	Ile	Gln	Trp	Ser		
			580					585					590				
Lys	Gly	Ala	Glu	Arg	Lys	Cys	Tyr	Arg	Pro	Leu	Arg	Pro	Cys	His	Cys		
		595					600					605					
His	Pro	Val	Leu	Ser	Glu	Arg	Glu	Ala	Val	Ile	Pro	Tyr	Ser	His	Arg		
	610					615					620						
Val	Phe	Asn	Thr	Ala	Asn	Pro	Leu	Lys	Gln	Gly	Glu	Cys	Gly	Leu	Arg		
625					630					635					640		

Ala	Ala	Ile	Ser	Asp	Met	Gly	Arg	Arg	Ile	Phe	Gln	Gln	Glu	Gln	Gln	
				645					650					655		
His	His	His	Arg	Ala	Tyr	Pro	Ala	Phe	Ala	Arg	Lys	Asn	Glu	Arg	His	
			660					665					670			
His	Ser	Glu	Ile	Tyr	Lys	Asn	Gly	Met	Gly	Gly	Trp	Leu	Asn	Lys	Ile	
		675					680					685				
Lys	Lys	Thr	Thr	Ile	Pro	Asn	Asn	Glu	Asn	Phe	Thr	Cys	Ile	Ser	Leu	
	690					695					700					
Gln	Leu	Leu	Trp	Gln	Leu	Tyr	Ser	Cys	Cys	Ile	Phe	Val	Gln	Ser	Glu	
705					710					715					720	
Gly	Asn	Leu	Gly	Ile	Gly	Ser	Val	Phe	Trp	Lys	Thr	Asn	Met	Thr	Ile	
				725					730					735		
Thr	Trp	Thr	Arg	Asn	Tyr	Ile	Asn	Ile	Pro	Tyr	Gly	Thr	Ile	Ile	Ser	
			740					745					750			
Leu	Phe	Met	Trp	Arg	Leu	Ser	Leu	Val	Phe	Leu	Phe	Tyr	Val	Ala	Ser	
		755					760					765				
Cys	Phe	Gln	Asn	Ser	Gln	Asn	Thr	Leu	Thr	Arg	Ile	Pro	Ala	Leu	Met	
	770					775					780					
Tyr	Leu	Phe	Arg	Thr	Lys	Ile	Asn	Lys	Leu	Ser	Phe	Leu	Arg	Lys	Trp	
785					790					795					800	
Met	Leu	Trp	Asn	Lys	Ser	Ser	Thr	His	Asn	Gly	Leu	Trp	Lys	Ser	Glu	
				805					810					815		
Ser	Arg	Met	Gln	Ser	Trp	Pro	Asn	Lys	Glu	Lys	Met	Thr	Leu	Leu	Cys	
			820					825					830			
Thr	Trp	Arg	Thr	Ile	Leu	Lys	Arg	Pro	Leu	His	Pro	Leu	Ser	Val	Ile	
		835					840					845				
Ala	Cys	Leu	Thr	Arg	Leu	Gln	Thr	Cys	Arg	Ile	Lys	Arg	Gln	Ser	Met	
	850					855					860					
Cys	Ile	Ser	Arg	Trp	Thr	Lys	Arg	Ile	Asp	Ser	Asn	Ser	Ser	Thr	Ser	
865					870					875					880	
Phe	Leu	Arg	Leu	His	Gly	Ile	Thr	Tyr	Lys	Arg	Arg	Gln	Lys	Arg	Thr	
				885					890					895		
Thr	Tyr	Thr	Ile	Cys	Trp	Cys	Arg	Pro	Met	Asn	Phe	Ile	Leu	Ser	Phe	
			900					905					910			
Pro	His	Met	Glu	Asn	Arg	Arg	Leu	Phe	Thr	Pro	Pro	Arg	Ile	Pro	Cys	
		915					920					925				
Pro	Ala	Thr	Leu	Ile	Asn	Ser	Arg	Glu	Ser	Leu	Thr	Thr	Phe	Lys	Thr	
	930					935						940				

Pro Leu His Thr Val Arg Ile Thr Ala Ser Leu Thr Leu Pro Arg Ala
 945 950 955 960
 Ser Pro Gly Met Trp Cys Gln Ser Asn Ser Arg Thr Leu Glu Ala Ser
 965 970 975
 Gln Lys Ile Ser Leu Pro Tyr Leu Lys Ser Ser Ile Gly Trp Thr Ile
 980 985 990
 Leu Val Leu Pro Ile Arg Val Ala Arg Asp Leu Asp Trp Arg Leu Gln
 995 1000 1005
 Lys Lys Leu Leu Phe Ser Met Glu Gly Arg Phe Thr Arg Lys Ala Met
 1010 1015 1020
 Ile Thr Ile Arg Arg Leu Gly Ser Phe Gln Arg Cys Gln Thr Trp Leu
 1025 1030 1035 1040
 Ile Lys Gly Gly Pro Lys Arg Cys Ile Phe Phe Arg Lys Ile Ser Arg
 1045 1050 1055
 Leu Ser Leu Leu Phe Leu Arg Lys Leu Thr Ile Tyr Glu Thr Ala Arg
 1060 1065 1070
 Ser Tyr Thr Val Asp Leu Ile Pro Glu Arg Ala Val Phe Val Leu Gln
 1075 1080 1085
 Arg Lys Ile Gln Asp Tyr His Trp His Pro Arg Phe Ile Trp Cys Leu
 1090 1095 1100
 Ser Gln Lys Gly Leu Val Leu Ile Met Asn Asn Ile Gly Ile Thr Val
 1105 1110 1115 1120
 Tyr Gly Cys Glu Gln Asp Glu Ala Asp Ala Phe His Ala Leu Ser Pro
 1125 1130 1135
 Arg Phe Gly Val Met Ala Thr Ile Ile Asn Ala Asn Val Ser Glu Ser
 1140 1145 1150
 Asn Ala Lys Ser Ala Pro Phe Asn Gln Cys Ile Ser Val Gly His Lys
 1155 1160 1165
 Ser Glu Ile Ser Ala Ser Ile Leu Leu Ala Leu Lys Arg Ala Gly Val
 1170 1175 1180
 Lys Tyr Ile Ser Thr Arg Ser Ile Gly Cys Asn His Ile Asp Thr Thr
 1185 1190 1195 1200
 Ala Ala Lys Arg Met Gly Ile Thr Val Asp Asn Val Ala Tyr Ser Pro
 1205 1210 1215
 Asp Ser Val Ala Asp Tyr Thr Met Met Leu Ile Leu Met Ala Val Arg
 1220 1225 1230
 Asn Val Lys Ser Ile Val Arg Ser Val Glu Lys His Asp Phe Arg Leu
 1235 1240 1245

Asp Ser Asp Arg Gly Lys Val Leu Ser Asp Met Thr Val Gly Val Val
 1250 1255 1260
 Gly Thr Gly Gln Ile Gly Lys Ala Val Ile Glu Arg Leu Arg Gly Phe
 1265 1270 1275 1280
 Gly Cys Lys Val Leu Ala Tyr Ser Arg Ser Arg Ser Ile Glu Val Asn
 1285 1290 1295
 Tyr Val Pro Phe Asp Glu Leu Leu Gln Asn Ser Asp Ile Val Thr Leu
 1300 1305 1310
 His Val Pro Leu Asn Thr Asp Thr His Tyr Ile Ile Ser His Glu Gln
 1315 1320 1325
 Ile Gln Arg Met Lys Gln Gly Ala Phe Leu Ile Asn Thr Gly Arg Gly
 1330 1335 1340
 Pro Leu Val Asp Thr Tyr Glu Leu Val Lys Ala Leu Glu Asn Gly Lys
 1345 1350 1355 1360
 Leu Gly Gly Ala Ala Leu Asp Val Leu Glu Gly Glu Glu Glu Phe Phe
 1365 1370 1375
 Tyr Ser Asp Cys Thr Gln Lys Pro Ile Asp Asn Gln Phe Leu Leu Lys
 1380 1385 1390
 Leu Gln Arg Met Pro Asn Val Ile Ile Thr Pro His Thr Ala Tyr Tyr
 1395 1400 1405
 Thr Glu Gln Ala Leu Arg Asp Thr Val Glu Lys Thr Ile Lys Asn Cys
 1410 1415 1420
 Leu Asp Phe Glu Arg Arg Gln Glu His Glu Asn Lys Ser Cys Asn Thr
 1425 1430 1435 1440
 Val Trp Gly Leu Leu Arg Gly Ala Arg Ile Gly Lys Ile Cys Asn Arg
 1445 1450 1455
 Asp Ser Arg His Arg Lys Ile Arg Ala Val Ile His Trp Asn Tyr Glu
 1460 1465 1470
 Ile Trp Cys Met Glu Asn Val Arg Lys Thr Leu Arg Gly Met Gly Lys
 1475 1480 1485
 Arg Gln Leu Leu Phe Ser Cys Thr Leu Ala Gly Lys Asn Ala Arg Ile
 1490 1495 1500
 Thr Cys Lys Glu Pro Ile Asn Gln Pro Cys Cys Ser Ile Phe Ser Phe
 1505 1510 1515 1520
 Ala Trp Gln Val Arg Arg Trp Ile His Thr Arg Ser Val Ile Val Arg
 1525 1530 1535
 Tyr Pro Phe Cys Arg Leu Arg Tyr Ser Lys Leu Ser Asn Leu Tyr Gly
 1540 1545 1550

Gln Ile Val Asp Ile His Arg Cys Glu Lys Cys Trp Asp Ser Tyr Ser	1555	1560	1565
Arg Leu Leu Gly Tyr Arg Ala Gly Gly Ser Tyr Val Tyr Leu Ser Cys	1570	1575	1580
Phe Cys Ala Gly Ala Phe Arg Leu Ile Leu Arg Cys Glu Lys Ser Gln	1585	1590	1595 1600
Arg Gly Arg Ile Gly Leu Arg Asn Ile Gly Lys Thr Ile Gln Gln Asn	1605	1610	1615
Leu Asn Ala Gly Cys Phe Gly Leu Gly Arg Leu Cys Gly Ile Gly Lys	1620	1625	1630
Gln Cys Arg Val Ser Cys Trp Arg Gly Gly Pro Asn Gln Ala Ala Val	1635	1640	1645
Arg Asn Leu Ser Tyr Ser Ser Gly Ser Arg Ala Gly Lys Arg Leu Lys	1650	1655	1660
Arg Ser Tyr Asn Arg Ser Arg Arg Pro Phe Ser Arg Gly Ala Arg Thr	1665	1670	1675 1680
Asp Thr Gly Asn Gly Lys Lys Asn Ile Ser Ala Arg Leu Arg Ser Ser	1685	1690	1695
Pro Cys Gly Tyr Val Phe Thr Arg Arg Pro His Cys Thr Glu Arg Ser	1700	1705	1710
Gln Tyr Ser Ala Arg Phe His Val Ile Gln Ser Leu Ser Pro Tyr Asp	1715	1720	1725
Gly Arg Cys Arg Tyr Cys Thr Ser Arg Thr Asp Pro Leu Asp Arg Ile	1730	1735	1740
Ser Val Lys Gly Val Ile Ser Met Glu Ile Gly Phe Thr Phe Leu Asp	1745	1750	1755 1760
Glu Ile Val His Gly Val Arg Trp Asp Ala Lys Tyr Ala Thr Trp Asp	1765	1770	1775
Asn Phe Thr Gly Lys Pro Val Asp Gly Tyr Glu Val Asn Arg Ile Val	1780	1785	1790
Gly Thr Tyr Glu Leu Ala Glu Ser Leu Leu Lys Ala Lys Glu Leu Ala	1795	1800	1805
Ala Thr Gln Gly Tyr Gly Leu Leu Leu Trp Asp Gly Tyr Arg Pro Lys	1810	1815	1820
Arg Ala Val Asn Cys Phe Met Gln Trp Ala Ala Gln Pro Glu Asn Asn	1825	1830	1835 1840
Leu Thr Lys Glu Ser Tyr Tyr Pro Asn Ile Asp Arg Thr Glu Met Ile	1845	1850	1855

Ser Lys Gly Tyr Val Ala Ser Lys Ser Ser His Ser Arg Gly Ser Ala
 1860 1865 1870
 Ile Asp Leu Thr Leu Tyr Arg Leu Asp Thr Gly Glu Leu Val Pro Met
 1875 1880 1885
 Gly Ser Arg Phe Asp Phe Met Asp Glu Arg Ser His His Ala Ala Asn
 1890 1895 1900
 Gly Ile Ser Cys Asn Glu Ala Gln Asn Arg Arg Arg Leu Arg Ser Ile
 1905 1910 1915 1920
 Met Glu Asn Ser Gly Phe Glu Ala Tyr Ser Leu Glu Trp Trp His Tyr
 1925 1930 1935
 Val Leu Arg Asp Glu Pro Tyr Pro Asn Ser Tyr Phe Asp Phe Pro Val
 1940 1945 1950
 Lys Thr Phe Asn Arg Cys Thr Asp Lys Leu Tyr Lys Leu Thr Leu Ser
 1955 1960 1965
 Ala Gly Asn Pro Thr Tyr Val Thr Gly Ser Gly Ile Tyr Ile Ile Val
 1970 1975 1980
 Leu Lys Met Gly Arg Ala Ile Leu Arg Ser Leu Ser Ala Cys Ala Ala
 1985 1990 1995 2000
 Ala Arg Pro Asp Asn Lys Thr Asp Arg Ile Glu Gly Trp Tyr Phe Thr
 2005 2010 2015
 Pro Pro Ile Val Asn Arg Gln Phe Ser Leu Val Lys Phe Ser Met Gly
 2020 2025 2030
 Ile Thr Tyr Glu Asn Ser Ser Thr Leu Val Ile Ile Val Asn Pro Val
 2035 2040 2045
 Gly Arg Asn Asn Leu Phe Thr Gly Gln Asn Gly Thr Ile Ser Asn Glu
 2050 2055 2060
 Ile Val Pro Phe Lys Gly Lys Ile Leu Glu Ile Phe His Thr Ser Asn
 2065 2070 2075 2080
 Tyr Ile Val Lys Glu Glu Thr Glu Asn Glu Glu Val Val Phe Phe Ile
 2085 2090 2095
 Val Ile Val Ile Leu Asn Ile Leu Arg Leu Leu Arg Ser Thr Val Phe
 2100 2105 2110
 Ser Gly Lys Ser Arg Ile Ser Lys Leu Ser Lys Ser Gln Arg Thr Phe
 2115 2120 2125
 Arg Lys Trp Asp Phe Lys Tyr Pro Arg Glu Asn Asn Tyr Arg Arg Thr
 2130 2135 2140
 Gly Leu Ser Arg Lys Ser Ala Ile Asn Gln Ile Ser Cys Ser Pro Arg
 2145 2150 2155 2160

[illegible]

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3190 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CAAAATATCA	CCTCATTTTT	GAGACAAGTC	TTATGAGACG	CTCTTAACTA	TGATTTTATC	60
AGTCTACTAC	ATTTGTATCA	ATAGAGTACA	CTCTATTGAT	ATATAATTGA	ACTAATAAAT	120
TGAAAATACA	GAAATGGAAT	GATACTGAAA	TGAAAATTGC	GAGAGGTAGA	GAATTGCTTA	180
CACCGGAACA	GAGACAGGCT	TTTATGCAA	TCCCTGAAGA	TGAATGGATA	CTGGGGACCT	240
ACTTCACTTT	TTCCAAACGG	GATTTAGAAA	TAGTTAATAA	GCGAAGGAGG	GAAGAAAACC	300
GTTTAGGATT	TGCTGTTCAA	TTAGCTGTTC	TTCGGTATCC	CGGTTGGCCA	TACACTCATA	360
TCAAAGCAT	CCCAGATTCG	GTCATACAAT	ATATATCGAA	ACAGATTGGT	GTTAGTCCAT	420
CCTCGCTTGA	TCATTATCCT	CAAAGGGAAA	ATACACTTTG	GGATCATTTG	AAAGAAATTC	480
GAAGTGAATA	CGACTTTGTA	ACTTTTACCC	TGAGTGAATA	TCGAATGACA	TTTAAGTACC	540

TTCATCAATT	AGCTTTGGAA	AATGGTGATG	CCATTCATCT	ACTGCATGAA	TGCATAGATT	600
TTCTAAGAAA	AAACAAAATT	ATACTGCCTG	CTATCACTAC	ACTTGAAAGA	ATGGTGTGGG	660
AAGCAAGGGC	AATGGCTGAA	AAGAAGCTAT	TTAATACGGT	TAGTAAATCT	CTAACAAATG	720
AGCAAAAAGA	AAAGCTTGAA	GGGATTATTA	CCTCGCAGCA	TCCATCCGAA	TCCAATAAAA	780
CGATATTGGG	TTGGTTAAAA	GAGCCACCGG	GTCATCCTTC	ACCCGAAACT	TTTCTAAAAA	840
TAATAGAACG	ACTCGAATAC	ATACGAGGAA	TGGATTTAGA	AACAGTGCAA	ATTAGTCATT	900
TGCACCGTAA	CCGCCTGTTG	CAGCTGTCTC	GCTTAGGCTC	AAGATACGAG	CCGTATGCAT	960
TCCGTGACTT	TCAAGAAAAT	AAACGTTATT	CGATATTAAC	CATCTATTTA	TTACAACCTA	1020
CTCAGGAGCT	AACGGATAAA	GCGTTTGAAA	TTCATGATAG	GCAAATACTT	AGTTTGTTAT	1080
CAAAAGGTCG	TAAGGCTCAA	GAGGAAATCC	AGAAACAAAA	CGGTAAAAAG	CTAAATGAGA	1140
AAGTTATACA	CTTTACGAAC	ATCGGACAAG	CATTAATTAA	AGCAAGAGAG	GAAAAATTAG	1200
ACGTTTTTAA	GGTTTTAGAA	TCGGTTATTG	AATGGAATAC	CTTTGTCTCT	TCAGTAGAAG	1260
AGGCTCAGGA	ACTTGCACGT	CCTGCCGACT	ATGATTATTT	AGACTTACTG	CAAAAACGGT	1320
TTTATTCACT	AAGAAAATAT	ACGCCAACGC	TATTAAGAGT	ATTGGAATTT	CATTCTACAA	1380
AGGCAAATGA	GCCACTTTTA	CAAGCTGTTG	AGATTATCCG	AGGAATGAAC	GAATCTGGAA	1440
AGCGAAAAGT	GCCTGATGAC	TCACCTGTGG	ATTTTATTTT	AAAACGATGG	AAAAGACATT	1500
TATACGAGGA	TGATGGTACA	ACAATTAATC	GTCATTACTA	TGAAATGGCT	GTTTTAACAG	1560
AACTTCGGGA	GCATGTTCCG	GCAGGAGATG	TTTCCATTGT	TGGCAGCAGA	CAATATAGGG	1620
ATTTTGAGGA	ATATTTGTTT	TCGGAAGATA	CATGGAATCA	ATCGAAGGGG	AATACGAGAT	1680
TATCAGTTAG	TTTATCATTC	GAAGATTATA	TAACGGAGAG	AACCAGCAGC	TTTAATGAAA	1740
GGTTAAAGTG	GTTAGCTGCC	AATTCCAATA	AGTTAGATGG	GGTTTCTCTT	GAAAAAGGAA	1800
AGCTATCACT	TGCACGCTTA	GAAAAAGATG	TTCCAGAAGA	AGCAAAAAAA	TTTAGTGCAA	1860
GCCTTTATCA	GATGCTACCA	AGAATAAAAT	TAAC TGATTT	ACTCATGGAT	GTGGCCCATA	1920
TAACAGGATT	TCATGAGCAA	TTCACTCATG	CTTCCAATAA	TCGAAAACCA	GATAAGGAAG	1980
AAACAATCAT	TATCATGGCT	GCCCTTTTAG	GAATGGGAAT	GAATATTGGC	TTGAGCAAGA	2040
TGGCCGAAGC	CACACCCGGA	CTTACATATA	AGCAACTAGC	CAATGTATCT	CAATGGCGCA	2100
TGTATGAAGA	TGCCATGAAT	AAAGCCCAAG	CCATATTAGT	AAACTTTCAT	CATAAATTAC	2160
AATTGCCTTT	CTATTGGGGC	GACGGTACAA	CATCTTCGTC	AGATGGTATG	AGAATGCAGC	2220
TAGGTGTTTC	ATCACTACAT	GCAGATGCAA	ATCCACATTA	TGGAACCTGGA	AAAGGAGCCA	2280

.CCATCTACCG	ATTTACAAGT	GATCAATTCT	CTTCTTACTA	CACAAAGATT	ATTCATACTA	2340
ATTCAGAGA	TGCGATTTCAT	GTTTTGGATG	GTTTGTTACA	TCATGAGACG	GATCTAAACA	2400
TAGAGGAACA	TTATACAGAC	ACTGCCGGTT	ACACTGACCA	AATATTCGGA	CTGACTCATT	2460
TATTAGGATT	TAAATTTGCC	CCAAGAATAA	GGGATTTATC	GGACTCAAAA	TTATTTACGA	2520
TAGATAAAGC	AAGTGAGTAT	CCAAAACTAG	AAGCCATTTT	ACGTGGACAA	ATAAATACAA	2580
AGGTCATTAA	AGAAAATTAT	GAGGATGTTT	TGCGATTAGC	TCATTCTATA	AGGGAGGGAA	2640
CAGTTTCAGC	ATCCCTTATT	ATGGGGAAGC	TAGGTTCCCTA	TTCAAGACAA	AACAGCTTAG	2700
CTACAGCCTT	ACGTGAGATG	GGCCGAATAG	AAAAAACGAT	CTTTATTTTG	AATTATATAT	2760
CGGATGAATC	ATTAAGAAGA	AAAATACAAA	GAGGATTGAA	TAAAGGAGAA	GCCATGAATG	2820
GATTGGCAAG	AGCTATTTTC	TTCGGAAAAC	AAGGTGAGCT	TAGAGAACGC	ACCATACAGC	2880
ATCAATTGCA	AAGAGCCAGT	GCTTTAAACA	TAATTATCAA	TGCTATAAGT	ATTTGGAATA	2940
CTCTCCACCT	AACAACAGCA	GTTGAATATA	AAAAACGGAC	AGGTAGCTTT	AATGAAGATT	3000
TGTTACACCA	TATGTCGCCC	TTAGGTTGGG	AACATATTAA	TTTACTAGGA	GAATACCATT	3060
TTAACTCAGA	GAAAGTAGTC	TCATTAAATT	CTTTAAGACC	ACTAAAACCTT	TCTTAACGTT	3120
GTAAAAACG	AGGGATTTCGT	CAGGAAAATA	GGCTTAGCGT	TGTAAATCCG	CATTTTCCTG	3180
ACGCTACCCC						3190

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1347 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GAGCTCTTCC	TTCAACGCAC	TTCTGTACCA	AGAGTTGTTG	TCCATTTGAT	CACTAACAAT	60
AGCTTCCCCT	GCTTTCTTCA	AGCCCTTTGT	CATAAAATCG	TTAGATTTTC	ATCATAAAAA	120
TACGAGAAAG	ACAACAGGAA	GACCGCAAAT	TTTCTTTTCT	TTTCCTAGGT	AACTGAATG	180
TAACCTTAAA	AGAAAAAAGG	AAAGGAAGAA	AATGATGAAA	AAAATTGCCG	TTTTATTTGG	240
AGGGAATTCT	CCAGAATACT	CAGTGTCACT	AACCTCAGCA	GCAAGTGTGA	TCCAAGCTAT	300
TGACCCGCTG	AAATATGAAG	TAATGACCAT	TGGCATCGCA	CCAACAATGG	ATTGGTATTG	360

GTATCAAGGA AACCTCGCGA ATGTTGCGAA TGATACTTGG CTAGAAGATC ACAAAAACCTG	420
TCACCAGCTG ACTTTTTTCTA GCCAAGGATT TATATTAGGA GAAAAACGAA TCGTCCCTGA	480
TGTCCTCTTT CCAGTCTTGC ATGGGAAGTA TGGCGAGGAT GGCTGTATCC AAGGACTGCT	540
TGAACTAATG AACCTGCCTT ATGTTGGTTG CCATGTCGCT GCCTCCGCAT TATGTATGAA	600
CAAATGGCTC TTGCATCAAC TTGCTGATAC CATGGGAATC GCTAGTGCTC CCACTTTGCT	660
TTTATCCCGC TATGAAAACG ATCCTGCCAC AATCGATCGT TTTATTCAAG ACCATGGATT	720
CCCGATCTTT ATCAAGCCGA ATGAAGCCGG TTCTTCAAAA GGGATCACAA AAGTAACTGA	780
CAAAACAGCG CTCCAATCTG CATTAACGAC TGCTTTTGCT TACGGTTCTA CTGTGTTGAT	840
CCAAAAGGCG ATAGCGGGTA TTGAAATTGG CTGCGGCATC TTAGGAAATG AGCAATTGAC	900
GATTGGTGCT TGTGATGCGA TTTCTCTTGT CGACGGTTTT TTTGATTTTG AAGAGAAATA	960
CCAATTAATC AGCGCCACGA TCACTGTCCC AGCACCATTG CCTCTCGCGC TTGAATCACA	1020
GATCAAGGAG CAGGCACAGC TGCTTTATCG AAACTTGGGA TTGACGGGTC TGGCTCGAAT	1080
CGATTTTTTC GTCACCAATC AAGGAGCGAT TTATTTAAAC GAAATCAACA CCATGCCGGG	1140
ATTTACTGGG CACTCCCGCT ACCCAGCTAT GATGGCGGAA GTCGGGTTAT CCTACGAAAT	1200
ATTAGTAGAG CAATTGATTG CACTGGCAGA GGAGGACAAA CGATGAACAC ATTACAATTG	1260
ATCAATAAAA ACCATCCATT GAAAAAAAT CAAGAGCCCC CGCACTTAGT GCTAGCTCCT	1320
TTTAGCGATC ACGATGTTTA CCTGCAG	1347

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 364 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met	Glu	Lys	Leu	Arg	Val	Gly	Ile	Val	Gly	Gly	Gly	Lys	Ser	Ala	Glu
1				5					10					15	
His	Glu	Val	Ser	Leu	Gln	Ser	Ala	Lys	Asn	Ile	Val	Asp	Ala	Ile	Asp
			20					25					30		
Lys	Ser	Arg	Phe	Asp	Val	Val	Leu	Leu	Gly	Ile	Asp	Lys	Gln	Gly	Gln
		35					40					45			
Trp	His	Val	Ser	Asp	Ala	Ser	Asn	Tyr	Leu	Leu	Asn	Ala	Asp	Asp	Pro
	50					55					60				

Ala	His	Ile	Ala	Leu	Arg	Pro	Ser	Ala	Thr	Ser	Leu	Ala	Gln	Val	Pro	65	70	75	80
Gly	Lys	His	Glu	His	Gln	Leu	Ile	Asp	Ala	Gln	Asn	Gly	Gln	Pro	Leu		85	90	95
Pro	Thr	Val	Asp	Val	Ile	Phe	Pro	Ile	Val	His	Gly	Thr	Leu	Gly	Glu	100	105		110
Asp	Gly	Ser	Leu	Gln	Gly	Met	Leu	Arg	Val	Ala	Asn	Leu	Pro	Phe	Val	115	120		125
Gly	Ser	Asp	Val	Leu	Ala	Ser	Ala	Ala	Cys	Met	Asp	Lys	Asp	Val	Thr	130	135	140	
Lys	Arg	Leu	Leu	Arg	Asp	Ala	Gly	Leu	Asn	Ile	Ala	Pro	Phe	Ile	Thr	145	150	155	160
Leu	Thr	Arg	Ala	Asn	Arg	His	Asn	Ile	Ser	Phe	Ala	Glu	Val	Glu	Ser		165	170	175
Lys	Leu	Gly	Leu	Pro	Leu	Phe	Val	Lys	Pro	Ala	Asn	Gln	Gly	Ser	Ser		180	185	190
Val	Gly	Val	Ser	Lys	Val	Thr	Ser	Glu	Glu	Gln	Tyr	Ala	Thr	Ala	Val	195	200		205
Ala	Leu	Ala	Phe	Glu	Phe	Asp	His	Lys	Val	Ile	Val	Glu	Gln	Gly	Ile	210	215		220
Lys	Gly	Arg	Glu	Ile	Glu	Cys	Ala	Val	Leu	Gly	Asn	Asp	Asn	Pro	Gln	225	230	235	240
Ala	Ser	Thr	Cys	Gly	Glu	Ile	Val	Leu	Thr	Ser	Asp	Phe	Tyr	Ala	Tyr		245	250	255
Asp	Thr	Lys	Tyr	Ile	Asp	Glu	Asp	Gly	Ala	Lys	Val	Val	Val	Pro	Ala	260	265		270
Ala	Ile	Ala	Pro	Glu	Ile	Asn	Asp	Lys	Ile	Arg	Ala	Ile	Ala	Val	Gln	275	280		285
Ala	Tyr	Gln	Thr	Leu	Gly	Cys	Ala	Gly	Met	Ala	Arg	Val	Asp	Val	Phe	290	295	300	
Leu	Thr	Pro	Glu	Asn	Glu	Val	Val	Ile	Asn	Glu	Ile	Asn	Thr	Leu	Pro	305	310	315	320
Gly	Phe	Thr	Asn	Ile	Ser	Met	Tyr	Pro	Lys	Leu	Trp	Gln	Ala	Ser	Gly		325	330	335
Leu	Gly	Tyr	Thr	Asp	Leu	Ile	Thr	Arg	Leu	Ile	Glu	Leu	Ala	Leu	Glu	340	345		350
Arg	His	Ala	Ala	Asn	Asn	Ala	Leu	Lys	Thr	Thr	Met					355	360		

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 306 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met	Thr	Asp	Lys	Ile	Ala	Val	Leu	Leu	Gly	Gly	Thr	Ser	Ala	Glu	Arg	
1				5					10					15		
Glu	Val	Ser	Leu	Asn	Ser	Gly	Ala	Ala	Val	Leu	Ala	Gly	Leu	Arg	Glu	
			20					25					30			
Gly	Gly	Ile	Asp	Ala	Tyr	Pro	Val	Asp	Pro	Lys	Glu	Val	Asp	Val	Thr	
		35					40					45				
Gln	Leu	Lys	Ser	Met	Gly	Phe	Gln	Lys	Val	Phe	Ile	Ala	Leu	His	Gly	
	50					55					60					
Arg	Gly	Gly	Glu	Asp	Gly	Thr	Leu	Gln	Gly	Met	Leu	Glu	Leu	Met	Gly	
65					70					75					80	
Leu	Pro	Tyr	Thr	Gly	Ser	Gly	Val	Met	Ala	Ser	Ala	Leu	Ser	Met	Asp	
				85					90					95		
Lys	Leu	Arg	Ser	Lys	Leu	Leu	Trp	Gln	Gly	Ala	Gly	Leu	Pro	Val	Ala	
			100					105						110		
Pro	Trp	Val	Ala	Leu	Thr	Arg	Ala	Glu	Phe	Glu	Lys	Gly	Leu	Ser	Asp	
		115					120					125				
Lys	Gln	Leu	Ala	Glu	Ile	Ser	Ala	Leu	Gly	Leu	Pro	Val	Ile	Val	Lys	
	130					135					140					
Pro	Ser	Arg	Glu	Gly	Ser	Ser	Val	Gly	Met	Ser	Lys	Val	Val	Ala	Glu	
145					150					155					160	
Asn	Ala	Leu	Gln	Asp	Ala	Leu	Arg	Leu	Ala	Phe	Gln	His	Asp	Glu	Glu	
				165					170					175		
Val	Leu	Ile	Glu	Lys	Trp	Leu	Ser	Gly	Pro	Glu	Phe	Thr	Val	Ala	Ile	
			180					185					190			
Leu	Gly	Glu	Glu	Ile	Leu	Pro	Ser	Ile	Arg	Ile	Gln	Pro	Ser	Gly	Thr	
		195					200					205				
Phe	Tyr	Asp	Tyr	Glu	Ala	Lys	Tyr	Leu	Ser	Asp	Glu	Thr	Gln	Tyr	Phe	
	210					215					220					
Cys	Pro	Ala	Gly	Leu	Glu	Ala	Ser	Gln	Glu	Ala	Asn	Leu	Gln	Ala	Leu	
225					230					235					240	

Val Leu Lys Ala Trp Thr Thr Leu Gly Cys Lys Gly Trp Gly Arg Ile
245 250 255

Asp Val Met Leu Asp Ser Asp Gly Gln Phe Tyr Leu Leu Glu Ala Asn
260 265 270

Thr Ser Pro Gly Met Thr Ser His Ser Leu Val Pro Met Ala Ala Arg
275 280 285

Gln Ala Gly Met Ser Phe Ser Gln Leu Val Val Arg Ile Leu Glu Leu
290 295 300

Ala Asp
305

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Leu Phe Pro Met Val Ile
1 5

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ser Thr Gln Asn Cys
1 5

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gly Glu Asp Gly Ser Ile Gln Gly
1 5

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Asn Thr Leu Pro Gly Phe Thr
1 5

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Gly Glu Asp Gly Thr Leu Gln Gly
1 5

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Asn Thr Ser Pro Gly Met Thr
1 5

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 259 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. faecium
- (B) STRAIN: BM4147

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Glu	Met	Asp	Val	Met	Glu	Gln	Lys	Leu	Asn	Thr	Leu	Lys	Arg	Thr	Leu	
1				5					10					15		
Glu	Lys	Arg	Glu	Gln	Asp	Ala	Lys	Leu	Ala	Glu	Gln	Arg	Lys	Asn	Asp	
			20					25					30			
Val	Val	Met	Tyr	Leu	Ala	His	Asp	Ile	Lys	Thr	Pro	Leu	Thr	Ser	Ile	
		35					40					45				
Ile	Gly	Tyr	Leu	Ser	Leu	Leu	Asp	Glu	Ala	Pro	Asp	Met	Pro	Val	Asp	
	50					55					60					
Gln	Lys	Ala	Lys	Tyr	Val	His	Ile	Thr	Leu	Asp	Lys	Ala	Tyr	Arg	Leu	
65					70					75					80	
Glu	Gln	Leu	Ile	Asp	Glu	Phe	Phe	Glu	Ile	Thr	Arg	Tyr	Asn	Leu	Gln	
				85					90					95		
Thr	Ile	Thr	Leu	Thr	Lys	Thr	His	Ile	Asp	Leu	Tyr	Tyr	Met	Leu	Val	
			100					105					110			
Gln	Met	Thr	Asp	Glu	Phe	Tyr	Pro	Gln	Leu	Ser	Ala	His	Gly	Lys	Gln	
		115					120					125				
Ala	Val	Ile	His	Ala	Pro	Glu	Asp	Leu	Thr	Val	Ser	Gly	Asp	Pro	Asp	
		130				135					140					
Lys	Leu	Ala	Arg	Val	Phe	Asn	Asn	Ile	Leu	Lys	Asn	Ala	Ala	Ala	Tyr	
145					150					155					160	
Ser	Glu	Asp	Asn	Ser	Ile	Ile	Asp	Ile	Thr	Ala	Gly	Leu	Ser	Gly	Asp	
			165						170					175		
Val	Val	Ser	Ile	Glu	Phe	Lys	Asn	Thr	Gly	Ser	Ile	Pro	Lys	Asp	Lys	
			180					185					190			
Leu	Ala	Ala	Ile	Phe	Glu	Lys	Phe	Tyr	Arg	Leu	Asp	Asn	Ala	Arg	Ser	
		195					200					205				

Ser Asp Thr Gly Gly Ala Gly Leu Gly Leu Ala Ile Ala Lys Glu Ile
210 215 220

Ile Val Gln His Gly Gly Gln Ile Tyr Ala Glu Ser Asn Asp Asn Tyr
225 230 235 240

Thr Thr Phe Arg Val Glu Leu Pro Ala Met Pro Asp Leu Val Asp Lys
245 250 255

Arg Arg Ser

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 256 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Escherichia coli

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Glu Ile Arg Val Met Pro Tyr Thr His Lys Gln Leu Leu Met Val Ala
1 5 10 15

Arg Asp Val Thr Gln Met His Gln Leu Glu Gly Ala Arg Arg Asn Phe
20 25 30

Phe Ala Asn Val Ser His Glu Leu Arg Thr Pro Leu Thr Val Leu Gln
35 40 45

Gly Tyr Leu Glu Met Met Asn Glu Gln Pro Leu Glu Gly Ala Val Arg
50 55 60

Glu Lys Ala Leu His Thr Met Arg Glu Gln Thr Gln Arg Met Glu Gly
65 70 75 80

Leu Val Lys Gln Leu Leu Thr Leu Ser Lys Ile Glu Ala Ala Pro Thr
85 90 95

His Leu Leu Asn Glu Lys Val Asp Val Pro Met Met Leu Arg Val Val
100 105 110

Glu Arg Glu Ala Gln Thr Leu Ser Gln Lys Lys Gln Thr Phe Thr Phe
115 120 125

Glu Ile Asp Asn Gly Leu Lys Val Ser Gly Asn Glu Asp Gln Leu Arg
130 135 140

Ser Ala Ile Ser Asn Leu Val Tyr Asn Ala Val Asn His Thr Pro Glu
145 150 155 160

Gly	Thr	His	Ile	Thr	Val	Arg	Trp	Gln	Arg	Val	Pro	His	Gly	Ala	Glu
				165					170					175	
Phe	Ser	Val	Glu	Asp	Asn	Gly	Pro	Gly	Ile	Ala	Pro	Glu	His	Ile	Pro
			180					185					190		
Arg	Leu	Thr	Glu	Arg	Phe	Tyr	Arg	Val	Asp	Lys	Ala	Arg	Ser	Arg	Gln
		195					200					205			
Thr	Gly	Gly	Ser	Gly	Leu	Gly	Leu	Ala	Ile	Val	Lys	His	Ala	Val	Asn
	210					215					220				
His	His	Glu	Ser	Arg	Leu	Asn	Ile	Glu	Ser	Thr	Val	Gly	Lys	Gly	Thr
225					230					235					240
Arg	Phe	Ser	Phe	Val	Ile	Pro	Glu	Arg	Leu	Ile	Ala	Lys	Asn	Ser	Asp
				245					250					255	

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 241 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Escherichia coli*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Ala Ser Glu Val Arg Ser Val Thr Arg Ala Phe Asn His Met Ala Ala
1 5 10 15

Gly Val Lys Gln Leu Ala Asp Asp Arg Thr Leu Leu Met Ala Gly Val
20 25 30

Ser His Asp Leu Arg Thr Pro Leu Thr Arg Ile Arg Leu Ala Thr Glu
35 40 45

Met Met Ser Glu Gln Asp Gly Tyr Leu Ala Glu Ser Ile Asn Lys Asp
50 55 60

Ile Glu Glu Cys Asn Ala Ile Ile Glu Gln Phe Ile Asp Tyr Leu Arg
65 70 75 80

Thr Gly Gln Glu Met Pro Met Glu Met Ala Asp Leu Asn Ala Val Leu
85 90 95

Gly Glu Val Ile Ala Ala Glu Ser Gly Tyr Glu Arg Glu Ile Glu Thr
100 105 110

Ala Leu Tyr Pro Gly Ser Ile Glu Val Lys Met His Pro Leu Ser Ile
115 120 125

Lys Arg Ala Val Ala Asn Met Val Val Asn Ala Ala Arg Tyr Gly Asn
130 135 140

Gly Trp Ile Lys Val Ser Ser Gly Thr Glu Pro Asn Arg Ala Trp Phe
145 150 155 160

Gln Val Glu Asp Asp Gly Pro Gly Ile Ala Pro Glu Gln Arg Lys His
165 170 175

Leu Phe Gln Pro Phe Val Arg Gly Asp Ser Ala Arg Thr Ile Ser Gly
180 185 190

Thr Gly Leu Gly Leu Ala Ile Val Gln Arg Ile Val Asp Asn His Asn
195 200 205

Gly Met Leu Glu Leu Gly Thr Ser Glu Arg Gly Gly Leu Ser Ile Arg
210 215 220

Ala Trp Leu Pro Val Pro Val Thr Arg Ala Gln Gly Thr Thr Lys Glu
225 230 235 240

Gly

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 231 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *E. faecium*
- (B) STRAIN: BM4147

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Met Ser Asp Lys Ile Leu Ile Val Asp Asp Glu His Glu Ile Ala Asp
1 5 10 15

Leu Val Glu Leu Tyr Leu Lys Asn Glu Asn Tyr Thr Val Phe Lys Tyr
20 25 30

Tyr Thr Ala Lys Glu Ala Leu Glu Cys Ile Asp Lys Ser Glu Ile Asp
35 40 45

Leu Ala Ile Leu Asp Ile Met Leu Pro Gly Thr Ser Gly Leu Thr Ile
50 55 60

Cys Gln Lys Ile Arg Asp Lys His Thr Tyr Pro Ile Ile Met Leu Thr
65 70 75 80

Gly	Lys	Asp	Thr	Glu	Val	Asp	Lys	Ile	Thr	Gly	Leu	Thr	Ile	Gly	Ala	
				85					90					95		
Asp	Asp	Tyr	Ile	Thr	Lys	Pro	Phe	Arg	Pro	Leu	Glu	Leu	Ile	Ala	Arg	
			100					105					110			
Val	Lys	Ala	Gln	Leu	Arg	Arg	Tyr	Lys	Lys	Phe	Ser	Gly	Val	Lys	Glu	
		115					120					125				
Gln	Asn	Glu	Asn	Val	Ile	Val	His	Ser	Gly	Leu	Val	Ile	Asn	Val	Asn	
	130					135					140					
Thr	His	Glu	Cys	Tyr	Leu	Asn	Glu	Lys	Gln	Leu	Ser	Leu	Thr	Pro	Thr	
145					150					155					160	
Glu	Phe	Ser	Ile	Leu	Arg	Ile	Leu	Cys	Glu	Asn	Lys	Gly	Asn	Val	Val	
				165					170					175		
Ser	Ser	Glu	Leu	Leu	Phe	His	Glu	Ile	Trp	Gly	Asp	Glu	Tyr	Phe	Ser	
			180					185					190			
Lys	Ser	Asn	Asn	Thr	Ile	Thr	Val	His	Ile	Arg	His	Leu	Arg	Glu	Lys	
		195					200					205				
Met	Asn	Asp	Thr	Ile	Asp	Asn	Pro	Lys	Tyr	Ile	Lys	Thr	Val	Trp	Gly	
	210					215					220					
Val	Gly	Tyr	Lys	Ile	Glu	Lys										
225					230											

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 239 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Met	Gln	Glu	Asn	Tyr	Lys	Ile	Leu	Val	Val	Asp	Asp	Asp	Met	Arg	Leu	
1				5					10					15		
Arg	Ala	Leu	Leu	Glu	Arg	Tyr	Leu	Thr	Glu	Gln	Gly	Phe	Gln	Val	Arg	
			20					25					30			
Ser	Val	Ala	Asn	Ala	Glu	Gln	Met	Asp	Arg	Leu	Leu	Thr	Arg	Glu	Ser	
		35					40					45				
Phe	His	Leu	Met	Val	Leu	Asp	Leu	Met	Leu	Pro	Gly	Glu	Asp	Gly	Leu	
	50					55					60					

Ser	Ile	Cys	Arg	Arg	Leu	Arg	Ser	Gln	Ser	Asn	Pro	Met	Pro	Ile	Ile	65	70	75	80
Met	Val	Thr	Ala	Lys	Gly	Glu	Glu	Val	Asp	Arg	Ile	Val	Gly	Leu	Glu	85	90	95	
Ile	Gly	Ala	Asp	Asp	Tyr	Ile	Pro	Lys	Pro	Phe	Asn	Pro	Arg	Glu	Leu	100	105	110	
Leu	Ala	Arg	Ile	Arg	Ala	Val	Leu	Arg	Arg	Gln	Ala	Asn	Glu	Leu	Pro	115	120	125	
Gly	Ala	Pro	Ser	Gln	Glu	Glu	Ala	Val	Ile	Ala	Phe	Gly	Lys	Phe	Lys	130	135	140	
Leu	Asn	Leu	Gly	Thr	Arg	Glu	Met	Phe	Arg	Glu	Asp	Glu	Pro	Met	Pro	145	150	155	160
Leu	Thr	Ser	Gly	Glu	Phe	Ala	Val	Leu	Lys	Ala	Leu	Val	Ser	His	Pro	165	170	175	
Arg	Glu	Pro	Ile	Ser	Arg	Asp	Lys	Leu	Met	Asn	Leu	Ala	Arg	Gly	Arg	180	185	190	
Glu	Tyr	Ser	Ala	Met	Glu	Arg	Ser	Ile	Asp	Val	Gln	Ile	Ser	Arg	Leu	195	200	205	
Arg	Arg	Met	Val	Glu	Glu	Asp	Pro	Ala	His	Pro	Arg	Tyr	Ile	Gln	Thr	210	215	220	
Val	Trp	Gly	Leu	Gly	Tyr	Val	Phe	Val	Pro	Asp	Gly	Ser	Lys	Ala	225	230	235		

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Met	Ala	Arg	Arg	Ile	Leu	Val	Val	Glu	Asp	Glu	Ala	Pro	Ile	Arg	Glu	1	5	10	15
Met	Val	Cys	Phe	Val	Leu	Glu	Gln	Asn	Gly	Phe	Gln	Pro	Val	Glu	Ala	20	25	30	
Glu	Asp	Tyr	Asp	Ser	Ala	Val	Asn	Gln	Leu	Asn	Glu	Pro	Trp	Pro	Asp	35	40	45	

Leu Ile Leu Leu Asp Trp Met Leu Pro Gly Gly Ser Gly Ile Gln Phe
 50 55 60
 Ile Lys His Leu Lys Arg Glu Ser Met Thr Arg Asp Ile Pro Val Val
 65 70 75 80
 Met Leu Thr Ala Arg Gly Glu Glu Glu Asp Arg Val Arg Gly Leu Glu
 85 90 95
 Thr Gly Ala Asp Asp Tyr Ile Thr Lys Pro Phe Ser Pro Lys Glu Leu
 100 105 110
 Val Ala Arg Ile Lys Ala Val Met Arg Arg Ile Ser Pro Met Ala Val
 115 120 125
 Glu Glu Val Ile Glu Met Gln Gly Leu Ser Leu Asp Pro Thr Ser His
 130 135 140
 Arg Val Met Ala Gly Glu Glu Pro Leu Glu Met Gly Pro Thr Glu Phe
 145 150 155 160
 Lys Leu Leu His Phe Phe Met Thr His Pro Glu Arg Val Tyr Ser Arg
 165 170 175
 Glu Gln Leu Leu Asn His Val Trp Gly Thr Asn Val Tyr Val Glu Asp
 180 185 190
 Arg Thr Val Asp Val His Ile Arg Arg Leu Arg Lys Ala Leu Glu Pro
 195 200 205
 Gly Gly His Asp Arg Met Val Gln Thr Val Arg Gly Thr Gly Tyr Arg
 210 215 220
 Phe Ser Thr Arg Phe
 225

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 129 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met Ala Asp Lys Glu Leu Lys Phe Leu Val Val Asp Asp Phe Ser Thr
 1 5 10 15
 Met Arg Arg Ile Val Arg Asn Leu Leu Lys Glu Leu Cys Phe Asn Asn
 20 25 30

Val	Glu	Glu	Ala	Glu	Asp	Gly	Val	Asp	Ala	Leu	Asn	Lys	Leu	Gln	Ala
	35						40					45			
Gly	Gly	Phe	Gly	Phe	Ile	Ile	Ser	Asp	Trp	Asn	Met	Pro	Asn	Met	Asp
	50					55					60				
Gly	Leu	Glu	Leu	Leu	Lys	Thr	Ile	Arg	Ala	Asp	Ser	Ala	Met	Ser	Ala
65					70					75					80
Leu	Pro	Val	Leu	Met	Val	Thr	Ala	Glu	Ala	Lys	Lys	Glu	Asn	Ile	Ile
				85					90					95	
Ala	Ala	Ala	Gln	Ala	Gly	Ala	Ser	Gly	Tyr	Val	Val	Lys	Pro	Phe	Thr
			100					105					110		
Ala	Ala	Thr	Leu	Glu	Glu	Lys	Leu	Asn	Lys	Ile	Phe	Glu	Lys	Leu	Gly
		115					120					125			

Met

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: E. faecium
 - (B) STRAIN: BM4147

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Met	Asn	Arg	Ile	Lys	Val	Ala	Ile	Leu
1			5					

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
 - (A) NAME/KEY: RBS
 - (B) LOCATION: 1..10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGAAAGGAGA

10

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

UCUUUCCUCC

10

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GCTGCAGATA AAAATTTAGG AGG

23

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CGCATGCTAT TATAAAAGCC AGTC

24

INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TGAAAGGGTG

10

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: rRNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: B. subtilis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGGGGGTTGG NNNNNNNNTT G

21

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: rRNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: B. subtilis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AGAACGAAAA NNNNNNATG

19